(19) World Intellectual Property Organization International Bureau





(43) International Publication Date 9 August 2001 (09.08.2001)

PCT

(10) International Publication Number WO 01/57188 A2

(51) International Patent Classification7:

C12N

English

(21) International Application Number: PCT/US01/03800

(22) International Filing Date: 5 February 2001 (05.02.2001)

(26) Publication Language: English

(30) Priority Data:

(25) Filing Language:

09/496,914 3 February 2000 (03.02.2000) US 09/560,875 27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 09/496,914 (CIP)
Filed on 3 February 2000 (03.02.2000)
US 09/560,875 (CIP)
Filed on 27 April 2000 (27.04.2000)

(71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San Jose, CA 95117 (US). **DRMANAC, Radoje, T.** [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

- (74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Coh, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

 without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

1/57188 A

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

2. BACKGROUND

5

10

15

20

25

30

35

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-1350. The polypeptides sequences are designated SEQ ID NO: 1351-2700. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, * corresponds to the stop codon.

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO:1-1350 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO:1-1350. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO:1-1350 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of SEQ ID NO:1-1350.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

5

10

15

20

25

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

5

10

15

20

25

30

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1 - 1350; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-1350. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing (e.g., SEQ ID NO: 1351-2700); (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-1350; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, e.g., in situ hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

5

10

15

20

25

30

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and form a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases o disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 DEFINITIONS

5

.. 10

15

20

25

30

35

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ

PCT/US01/03800 WO 01/57188

cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

10

15

20

25

30

35

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can

be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-1350.

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-1350. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4²⁰ possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match $(1 \div 4^{25})$ times the increased probability for mismatch at each nucleotide position (3×25) . The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

5

10

15

20

25

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements *e.g.* repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

5

10

15

20

25

30

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, e.g., recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

5

10

15

20

25

30

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

5

10

15

20

25

30

35

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, e.g., polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., E. coli, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134-143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (*i.e.*, hybridization to filter-bound DNA in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (*i.e.*, washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

5

.10

15

20

25

30

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

5

10

15

20

25

30

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (i.e., the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, e.g., mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (e.g., via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

20

25

30

35

5

10

15

4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO:1-1350; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO:1351-2700; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO:1351-2700. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO:1-1350; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 1351-2700. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic

domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

5

10

15

20

25

30

35

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO:1-1350 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO:1-1350 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO:1-1350 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, e.g., at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO:1-1350, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that

are selective for (i.e. specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO:1-1350, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO:1-1350 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO:1-1350, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

5

10

15

20

25

30

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, e.g., by substituting first with conservative choices (e.g., hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (e.g., hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

10

15

20

25

30

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., DNA 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, Nucleic Acids Res. 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., Gene 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and Current Protocols in Molecular Biology, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO:1-1350, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are

5

10 ·

15

20

25

30

known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

5

10

15

20

25

30

35

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

4.3 ANTISENSE

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1-1350, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

5

10

15

20

25

NO:1351-2700 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO:1-1350 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

10

15

20

Given the coding strand sequences encoding a nucleic acid disclosed herein (e.g., SEQ ID NO:1-1350), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an -a nomeric nucleic acid molecule. An -a nomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual -units, the strands run parallel to each other (Gaultier et al. (1987) Nucleic Acids Res 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue et al. (1987) FEBS Lett 215: 327-330).

4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

5

10

15

20

25

30

designed based upon the nucleotide sequence of a DNA disclosed herein (i.e., SEQ ID NO:1-1350). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, e.g., Cech et al. U.S. Pat. No. 4,987,071; and Cech et al. U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., (1993) Science 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991)

Anticancer Drug Des. 6: 569-84; Helene. et al. (1992) Ann. N.Y. Acad. Sci. 660:27-36; and Maher (1992) Bioassays 14: 807-15.

-10

. 15

20

25

30

35

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) Bioorg Med Chem 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. (1996) above; Perry-O'Keefe et al. (1996) PNAS 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup et al. (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn et al. (1996) Nucl Acids Res 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag et al. (1989) Nucl Acid Res 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen et al. (1975) Bioorg Med Chem Lett 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

25

30

35

5

10

15

20

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in coamplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces strains, Candida, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include Escherichia coli, Bacillus subtilis, Salmonella typhimurium, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

5

10

15

20

25

30

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

5

10

15

20

25

30

35

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO:1351-2700 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:1-1350 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO:1-1350 or (b)

polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO:1351-2700 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO:1351-2700 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO:1351-2700.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

5

10

15

20

25

30

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, Protein Purification: Principles and Practice, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: A Laboratory Manual; Ausubel et al., Current Protocols in Molecular Biology. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for *e.g.*, small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO:1351-2700.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

5

10

15

20

25

30

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBatTM kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

5

10

15

20

25

30

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearlTM or Cibacrom blue 3GA SepharoseTM; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, *e.g.*, targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, *e.g.*, antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

15

20

25

30

35

10

5

4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

10

15

20

25

30

35

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

4.8 GENE THERAPY

5

10

15

20

25

30

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected ex vivo, in situ, or in vivo by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered in vivo to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.9 TRANSGENIC ANIMALS

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

5

10

15

20

25

30

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

5

10

15

20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

4.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

4.10.1 RESEARCH USES AND UTILITIES

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

5

10

15

20

25

30

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

4.10.2 NUTRITIONAL USES

5

10

15

20

25

30

35

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells 20 include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology, J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto, 1991: deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse 25 and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. 30 J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober,

35

5

10

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

10

20

25

30

35

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells in vivo or ex vivo is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce 15 large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

5

10

15

20

25

30

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support *e.g.* as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) · as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

15

10

5

4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

20

25

30

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

5

10

15

20

25

30

35

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine,

kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

20

25

30

5

10

15

4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (e.g., anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on aflergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastbom et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

10

15

20

25

30

35

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue

transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial

5

10

15

20

25

30

immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

5

10

15

20

25

30

35

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J.

Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

5

10

15

20

25

30

4.10.8 ACTIVIN/INHIBIN ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polypucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

5

10

15

20

25

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

10

15

20

25

30

35

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine.

Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These in vitro models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions

5

10

15

20

25

30

and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14. Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

30

35

10

15

20

25

4.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science* 282:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., Mol. Biotechnol, 9(3):205-23 (1998); Hruby et al., Curr Opin Chem Biol, 1(1):114-19 (1997); Dorner et al., Bioorg Med Chem, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

5

10

15

20

25

30

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (i.e., increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

5

10

15

20

25

4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

4.10.16 LEUKEMIAS

Leukemias and related disorders may be treated or prevented by administration of a
therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the
invention. Such leukemias and related disorders include but are not limited to acute leukemia,
acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic,
myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic
(granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see
Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of

35

5

10

15

therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

- (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

10

15

20

- (iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;
- (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;
- (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;
- (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis:
- (vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and
 - (viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

- (i) increased survival time of neurons in culture;
- (ii) increased sprouting of neurons in culture or in vivo;
- (iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or
 - (iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape

35 (such as, for example, breast augmentation or diminution, change in bone form or shape);

5

10

15

20

effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

4.10.19 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or

absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

4.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

4.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

4.11.1 EXAMPLE

5

10

15

20

25

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01µg/kg to 100 mg/kg of body weight, with the preferred dose being about $0.1 \mu g/kg$ to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

5

10

15

20

25

30

35

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth

factor (PDGF), transforming growth factors (TGF- α and TGF- β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other

5

10

15

20

25

30

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

4.12.1 ROUTES OF ADMINISTRATION

10

15

20

25

30

35

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers

comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

5

10

15

20

25

30

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, tale, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

10

15

20

25

30

35

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral

administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other

5

10

15

20

25

30

sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

5

10

15

20

25

30

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically

acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1 µg to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

5

10

15

20

25

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

5

10

15

20

25

30

35

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications.

Particularly domestic animals and thoroughbred horses, in addition to humans, are desired

patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

20

25

30

5

10

15

4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC₅₀ as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD50 (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD₅₀ and ED₅₀. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from in vitro data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 μ g/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 μ g/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

4.12.4 PACKAGING

10

15

20

25

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

4.13 ANTIBODIES

5

10

15

20

25

30

35

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} , F_{ab} and $F_{(ab)2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG_1 , IgG_2 , and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, (for example the amino acid sequence shown in SEQ ID NO: 1351), and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, e.g., a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will

indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, e.g., Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

5.13.1 Polyclonal Antibodies

5

10

15

20

25

30

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

5.13.2 Monoclonal Antibodies

5

10

15

20

25

30

35

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, <u>J. Immunol., 133</u>:3001 (1984); Brodeur et al., <u>Monoclonal Antibody Production Techniques and Applications</u>, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

5

10

15

20

25

30

35

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, <u>Anal. Biochem.</u>, <u>107</u>:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium.

Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for

example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

5.13.2 Humanized Antibodies

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')2 or other antigenbinding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

5.13.3 Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

10

15

20

25

30

Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al. (Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the XenomouseTM as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the

immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

5.13.4 Fab Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab)2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab)2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_{v} fragments.

5.13.5 Bispecific Antibodies

5

10

15

20

25

30

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

5

10

15

20

25

30

35

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')₂ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure

wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., <u>J. Immunol.</u> 147:60 (1991). Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on

5

10

15

20

25

30

a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc R), such as Fc RI (CD64), Fc RII (CD32) and Fc RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

5.13.6 Heteroconjugate Antibodies

5

10

15

20

25

30

35

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced antitumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of

bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ²¹²Bi, ¹³¹I, ¹³¹In, ⁹⁰Y, and ¹⁸⁶Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

4.14 COMPUTER READABLE SEQUENCES

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled

5

10

15

20

artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO:1-1350 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO:1-1350 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored

therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

4.15 TRIPLE HELIX FORMATION

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA.

5

10

15

20

25

30

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem.

5 6:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

4.16 DIAGNOSTIC ASSAYS AND KITS

15

20

25

30

35

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary.

Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization,

amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

4.17 MEDICAL IMAGING

5

10

15

20

25

30

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide in vivo at the target site.

4.18 SCREENING ASSAYS

10

15

20

25

30

35

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:1-1350, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
 - (b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polypucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

5

10

15

20

25

30

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

4.19 USE OF NUCLEIC ACIDS AS PROBES

5

10

15

20

25

30

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO:1-1350. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO:1-1350 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, *i.e.*, small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata et al., 1985; Dahlen et al., 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller et al., 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, *e.g.*, Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen et al., (1991) Anal. Biochem. 198(1) 138-42).

5

10

15

20

25

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

5

10

15

20

25

30

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm₇), is then added to a final concentration of 10 mM 1-MeIm₇. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm₇, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, *e.g.*, Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, CviJI, described by Fitzgerald et al. (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

5

10

15

20

25

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease CviII normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme ($CviII^{**}$), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a $CviII^{**}$ digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that $CviII^{**}$ restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

4.22 PREPARATION OF DNA ARRAYS

5

10

15

20

25

30

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm², depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm² and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

5.0 EXAMPLES

5.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems

5

10

15

20

25

(ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

5.2 EXAMPLE 2

Novel Contigs

5

10

15

20

25

30

The novel contigs of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The sequences for the resulting nucleic acid contigs are designated as SEQ ID NO: 1-1350 and are provided in the attached Sequence Listing. The contigs were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Table 3 sets forth the novel predicted polypeptides (including proteins) encoded by the novel polynucleotides (SEQ ID NO:189-282) of the present invention, and their corresponding nucleotide locations to each of SEQ ID NO: 189-282. Table 3 also indicates the method by which the polypeptide was predicted. Method A refers to a polypeptide obtained by using a software program called FASTY (available from http://fasta.bioch.virginia.edu) which selects a polypeptide based on a comparison of the translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference). Method B refers to a polypeptide obtained by using a software program called GenScan for human/vertebrate sequences (available from Stanford University, Office of Technology Licensing) that predicts the polypeptide based on a probabilistic model of gene structure/compositional properties (C. Burge and S. Karlin, J. Mol. Biol., 268:78-94 (1997), incorporated herein by reference). Method C refers to a polypeptide obtained by using a Hyseq proprietary software program that translates the novel polynucleotide and its complementary strand into six possible amino acid sequences (forward and reverse frames) and chooses the polypeptide with the longest open reading frame.

The nearest neighbor results for SEQ ID NO: 1-1350 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq database October 12, 2000, update 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the

closest homologue for SEQ ID NO:1-1350. The nearest neighbor results for SEQ ID NO: 1-1350 are shown in Table 2 below.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NO: 1-1350. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/496,914.

TABLE 1

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
adult brain	GIBCO	AB3001	111 151 188 215 662-665 877 910 927
	<u> </u>		976 1233 1319
adult brain	GIBCO	ABD003	41 49 74 101 111 120 132 141-142 151
	1		217 225 238 271 317 404 446 469 503
	j		513-514 535 550 564 573 666-669 798
			898 910 927 976 1067 1083 1085 1178
			1254
adult brain	Clontech	ABR001	39 216 238 327 356 535 927 1056 1121
			1178-1180 1199 1251
adult brain	Clontech	ABR006	74 611 949 1034 1136
adult brain	Clontech	ABR008	14 32 41 61 81 86 89 120 132 138 145
	,	,	147 188 197 208 225 227-239 250 300-
			303 312 316 328-331 340 357-362 374
		,	380 384-391 408 414 446 448 464-467
			483 488 495-496 505 512 521 535 550
			566 571 577 585 590 594 598 634 641
			658 666 683 725 742 764 767 786 801
		ł	805 810 823 826 829 831 836 841 887-
			923 927 934 943 950-951 963 976 995
		1	1000-1001 1006 1026 1034 1048 1057-
			1067 1086 1088 1090 1118 1120 1122-
			1128 1142 1162 1181-1192 1199 1204 1218-1219 1225 1232 1253 1267 1271-
3.1.1	Glto-b	ADDOLL	1306 1342 1347 1349-1350 49 238 1219
adult brain	Clontech	ABR011	74 238
adult brain	BioChain	ABR012	868 1268
adult brain	Invitrogen	ABR013	49 117 138 191 217 252 291 305 535
adult brain	Invitrogen	ABT004	
			566 596 663 670 746 798 816-819 876 892 898 922 943 963 1034-1036 1121
cultured	Charles	ADP001	41 74 101 138 211 238 304 537 582
preadipocytes	Strategene	ADPOUL	740 798 883 943 976 1067
adrenal gland	Clontech	ADR002	49 74 101 111 120 127 151 215 238
acrenal gland	Ciontech	ADROUZ	240-247 316 330 363-364 404 414 534-
			535 833 924-940 950 963 976 1001
			1003 1067-1070 1118 1156 1193-1200
			1325
adult heart	GIBCO	AHR001	38 49 71-72 74-77 79 92 99 101 111
adult hourt	OIDCO	7111111001	118 129 132 138 151 158-163 182 195-
			203 215 217 238 264 269 353 384 398
	ŀ		408 434-439 446 504 512-513 519 537
			562-573 577 611-614 616-619 658 661
	Ì		671-672 722 734 757-773 815 828-835
			874 891 898 919 926-927 976 988
			1021 1037 1041 1062 1067 1071 1080
			1083 1093 1122 1131 1185 1201 1254
			1308 1331 1335
adult kidney	GIBCO	AKD001	41 49 51 71-74 78-85 94 100-101 103-
			107 111 119-120 138 151 157 215 217-
			218 238 250 264 294 304 384 404 440
		*	446 454 477 504-505 509 514 518-519
			535 537 564 574-583 620-627 639 653
			673-675 705 753 789 831 844 851 859
1	3		1
			877 909 918 927 956 963 976 1067
·			877 909 918 927 956 963 976 1067 1074 1083 1095 1178 1302 1331 1335
adult kidnev	Invitrogen	AKT002	1
adult kidney	Invitrogen	AKT002	1074 1083 1095 1178 1302 1331 1335
adult kidney	Invitrogen	AKT002	1074 1083 1095 1178 1302 1331 1335 11-12 41 49 111-112 215-217 294 316

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			518 537 545 549 580 582 592 594 634
			640 651-652 676-678 725 851 873 918
			952 976 1042 1067 1076 1083 1152
lymph node	Clontech	ALN001	8 111 121 151 180-182 188 215 537
	Ì		545 549 651 679-682 789 804-810 868
			873 927 952 976 1042 1059 1335
young liver	GIBCO	ALV001	8 64 79 111 186 215-216 238 446 514
			519 537 564 653 683-684 698 753 798
			813 833 840 858 927 976 1038-1039
			1051 1085 1224 1245 1256
adult liver	Invitrogen	ALV002	40 71 292-293 305 384 468-469 496
			505 657 675 714 753 832 844 941-942
	L		976 1040 1076 1256 1293
adult liver	Clontech	ALV003	976
adult ovary	Invitrogen	AOV001	8 32 36 38 41 49 51 71 74 79-80 101
			104 111 120 122-125 138 140 143-149
			151 188-190 207-212 215-217 238 264
			316 384 409 440 445-446 496 504 512
			514 518-519 535 537 549-550 564 566
			571 580 582 600 618 638 657 667 681
			685-697 699 705 722 735-744 761 771
			815 833 842-865 868 875-876 918 926-
			927 950 952 963 976 1023 1042 1048
			1051 1059 1072 1076 1083 1117 1120
			1124 1131 1144 1174 1224 1268 1331
			1335
adult placenta	Clontech	APL001	102 217 238 537 641 700
placenta	Invitrogen	APL002	663 851 1048
adult spleen	GIBCO	ASP001	8 45 74 111 132 140 151 185 217 238
			294 414 446 477 504 514 534 545 549
			592 722 873 883 952 976 1041-1042
			1083 1093-1094 1152 1224
testis	GIBCO	ATS001	72 107 111 113 126 140 151 183 215
			238 446 497 537 642 701-706 811 877
			927 962 976 1083 1117 1131
adult bladder	Invitrogen	BLD001	41 151 191 402-405 409 414 496 545
			592 607 706 873 952 1178 1329-1335
bone marrow	Clontech	BMD001	8 58-62 65-68 74 79 108 111 116 137
			147 151 164-174 213-215 238 305-307
			374 404 446 460 466 516 519 534 538-
			541 544-546 549-554 566 584 586 592
			596 607 610 628-629 643-645 652 707-
	j		708 774-789 844 866-871 873 919 927
u			952 963 976 998 1034 1042 1064 1083
			1085 1120 1132 1152 1225 1229 1268
			1307 1310
bone marrow	Clontech	BMD002	6 8 37-38 52 74 77 105 111 129 132
			210 317 510-511 545 549 581 598 628
			638 724 766 789 844 860 868 873 919
			927 952 963 968 976 1042 1111 1141
			1160-1161 1229 1266 1346
bone marrow	Clontech	BMD004	111 238 282 549 1083
adult colon	Invitrogen	CLN001	52 260 264 299 494 536 545 564 592
	ļ		844 873 877 952 976 1042 1152 1268
·			1336-1337
adult cervix	BioChain	CVX001	49 51 129 132 151 205 207 238 332-
			335 365-367 392-401 440 466 470-471
	(518 537 597 629 832 877 927 976 1006
		1	1 1005 1117 1100 1104 1100 1000 1005
I			1085 1117 1129-1134 1192 1202-1205
diaphragm	BioChain	DIA002	1085 1117 1129-1134 1192 1202-1205 1219 1309-1328 74 976 1083

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
endothelial cells	Strategene	EDT001	32 40-41 49 74 79 101 111 120 132
	J		138 151 204-206 215-217 238 269 316
			414 433 505 510 513 550 555 580 582
	1		596 675 722 745 798 814 836-841 851
	į.		918 976 1041 1043 1073 1083 1131
			1331
Genomic clones	Genomic DNA	EPM001	525-532 927
from the short arm	from Genetic)	
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM003	47 525
from the short arm	from Genetic		1
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM004	525 927
from the short arm	from Genetic		
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM005	531
from the short arm	from Genetic	ļ	
of chromosome 8	Research		
esophagus	BioChain	ESO002	74 138 238
fetal brain	Clontech	FBR001	441-442 927
fetal brain	Clontech	FBR004	215 893 927 1001
fetal brain	Clontech	FBR006	48 61 101 120 132 138 140 147 208
}			225 271 317 319 336 359 368 405-414
	1		519 550 571 594 686 715 722 764 824
	Ì		829 836 859 909 927 943 947 963 1057
)	1067-1068 1104 1135-1140 1162 1206- 1207 1235 1268 1288 1307-1308 1319
			1338-1350
fetal brain	Clontech	FBRs03	111 446
fetal brain	Invitrogen	FBT002	41 51 120 151 192-194 264 504 512
iciai viani	mvinogen	FB1002	535 683 761 798 820-827 844 876 909
	1	}	963 976 1026 1048 1083 1144 1302
fetal heart	Invitrogen	FHR001	446 566 761
fetal kidney	Clontech	FKD001	51 74 111 127 140 151 184 294 537
loui Ridirey	Cionicon	112001	550 630-631 1319
fetal kidney	Clontech	FKD002	111 976 1083
fetal kidney	Invitrogen	FKD007	238 974
fetal lung	Clontech	FLG001	463 566 976 1074 1083 1093
fetal lung	Invitrogen	FLG003	41 238 330 407 415-416 537 573 844
			859 1048 1083 1116 1192
fetal liver-spleen	Columbia	FLS001	8 14 34-35 37 41 43 49 51 54-56 63-64
ł	University		69-71 74 77 79 87-90 101 107 110-111
			114 120 128-131 138 140 147 150-155
			197 210 215 217 225 238 312 367 384
		}	414 440 446 460 468 483 496 504-507
	1		511-515 518-519 523 533-535 537 541
	1		544-545 547-550 555-560 564 566 571
		}	577 582 585-586 598 636 646-647 649
1			652 664 698 709-710 714 722-723 731
	}		735-736 746-753 761 784 798 823 829
	1		832 844 851 858-859 868 873 876 898
		1	927 943 949 952 963 976 984 1002 1021 1023 1040 1042 1044 1050 1083
	1		1021 1023 1040 1042 1044 1050 1083
1			1217 1251 1254 1256 1302 1308 1311
	}	}	1217 1251 1254 1256 1302 1308 1311
fotal liver calcon	Columbia	EI SOO2	8 36-37 41-46 49 54 64 71 74 79 101
fetal liver-spleen	Columbia	FLS002	f
	University		111 120 129 147 207 210 215-216 238 250 330 353 359 366 383-384 414 478
	1		505 508-509 511 515-524 534-535 537
			544-545 564 566 571 577 591 598 638

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			663 671 698 714 722 725 727 751 798
			851 859 873 876 909 927 949 952 983-
			984 1002 1023 1042-1044 1085 1095
			1131 1144 1178 1199 1233 1240-1270
			1331 1340
fetal liver-spleen	Columbia University	FLS003	64 535 976 1256
fetal liver	Invitrogen	FLV001	8 101 120 138 217 446 468 535 566
		12,001	580 722 730 749 844 918 943 976 1051
	İ		1256 1331
fetal liver	Clontech	FLV004	537 926 1256
fetal muscle	Invitrogen	FMS001	51 111 264 312 369-370 404 417-421
			425 535 537 577 598 614 836 857 1141
		1	1208 1268
fetal muscle	Invitrogen	FMS002	537
fetal skin	Invitrogen	FSK001	13-26 32 41 51 89 107 111 147 151
		j	225 264 316 405 422-429 488-494 496
			519 534-535 537 566 675 732 859 876-
			877 898 947 949-950 963 976 1001
			1062 1076 1083 1117 1144 1165 1268
			1281
fetal skin	Invitrogen	FSK002	537 812
fetal spleen	BioChain	FSP001	87 549
umbilical cord	BioChain	FUC001	27-33 41 49 151 215 238 248-249 301
			316 446 495-503 519 521 534-535 537
			582 634 691 877 883 927 944-950 963
			976 1001 1075 1142-1143 1171 1218
			1243 1308
fetal brain	GIBCO	HFB001	41 49 57 79 87 103 111 120 132-135
			138 145 151 188 197 207 215 238 264
			271 294 316 367 414 440 446 466 504
			513-514 535 542-543 550 564 571 596
			635 648-654 675 711-715 722-723 798
	1		832 872 876 883 927 976 1095 1144
		<u> </u>	1168 1171 1178 1211 1335
macrophage	Invitrogen	HMP001	238
infant brain	Columbia	IB2002	49-50 77 81 89 105 111 136-138 140
	University		151 161 175-179 185 216-217 264 295
			299 308-310 371-373 462 476 504 511-
	- {		513 533 537 564 566 571 655-657 662
			683 716-720 723 752 790-803 829 832
			858-859 876 898 909 949 976 1045-
			1047 1076-1087 1090 1093 1116 1122
			1144 1209-1213 1225 1233 1256 1319
:	101 1:	770000	1341
infant brain	Columbia	IB2003	41 50 77 104 132 215 238 508 512-513
	University		519 566 655 714 794 918 943 976 1067
:-C	1001	HD) (000	1092-1093 1233
infant brain	Columbia	IBM002	311 472-473 753 1214
	University	<u> </u>	
infant brain	Columbia	IBS001	51 111 376 474 790 876 949 1144 1204
lung Charles	University	I FDoo:	1221
lung, fibroblast	Strategene	LFB001	151 316 462 514 534 582 675 939 1131
lung tumor	Invitrogen	LGT002	1-7 41 74 79 94 115 120 138-139 156
	[215 217 269 280 296 337 374-375 384
	İ		404 446 454 475-480 498 514 518-519
		1	522 537 545 564 577 597 653 658 705
	1		721-724 754-756 779 859 868 872-874
			876-877 919 927 949 951-952 959 976
			1002 1042 1048-1053 1076 1083 1088-
		1	1089 1131 1144-1147 1216-1218 1229

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
I issue Origin	KNA Source	Hysed Library Name	1293 1311
lymphocytes	ATCC	LPC001	41 74 111 132 151 253 316 446 550
Гупгриосуюз	11100	LI COUL	634 844 927 976 1085 1268
leukocyte	GIBCO	LUC001	8 11 41 74 86 91-98 101 109 111 120
louriou) to			147 151 212 215 218 238 252 288 312-
	1		314.316 338 359 408 427 443-447 505
			510 512 514 518 534 545 549-550 561
			564 566 571 577 580 582 587-609 615
			632-638 658-659 698 714 725-728 832
			836 841 859 866 873-874 882-883 918-
			919 927 943 952 963 976 1042 1076
			1083 1090 1148 1152 1168 1195 1219-
			1220 1224
leukocyte	Clontech	LUC003	74 100 215 232 238 339-341 446 545
			657 660 729 873 883 927 952 963 1008
			1042 1116 1120 1149-1150 1215 1222
Melanoma from cell	Clontech	MEL004	210 215 238 342 534 545 592 722 873
line ATCC #CRL	[919 929 939 952 976 1071 1118 1218
1424	<u> </u>		1235 1245
mammary gland	Invitrogen	MMG001	8-10 40-41 49 73 80 114 138-140 147
			217 250-256 264 297-299 305 377-378
	ì		398 446 481-486 505 512 537 545 549
	Į.		571 592 725 730-733 816 829 836 844
	[868 873 876-877 898 926 943 951-960
	ļ		963 976 995 1034 1042 1048 1054- 1055 1076 1083 1091 1093 1116-1117
	ĺ		
	0	37770001	1124 1152 1302 39 101 111 138 238 361 1225 1251
induced neuron cells	Strategene	NTD001	1319
retinoid acid induce I	Strategene	NTR001	74 225 976
neuronal cells	Sualegene	MIROUI	14 223 310
neuronal cells	Strategene	NTU001	129 225 238 304 313 361 657 976
pituitary gland	Clontech	PIT004	976
placenta	Clontech	PLA003	38 976
prostate	Clontech	PRT001	111 188 238 257-258 564 724 961-966
•			1067 1095
rectum	Invitrogen	REC001	238 430-431 841 859 868 963 1001
			1116
salivary gland	Clontech	SAL001	8 151 402 432-433 446 496 868 952
			976 1083 1120 1151 1184
small intestine	Clontech	SIN001	8 101 147 215 259-266 446 462 505
			545 592 660 789 836 866 873 927 952
•			963 967-978 1042 1120 1152 1223-
			1224
skeletal muscle	Clontech	SKM001	238 302 927 943 992 1031
spinal cord	Clontech	SPC001	74 111 132 151 215-216 238 264 267-
			270 343-344 353 379 516 537 566 740
			828 927 976 979-994 1092 1153-1159
	 	dr. c.	1225 1250
adult spleen	Clontech	SPLc01	698 859 1042 210 238 271-272 537 580 705 918 952
stomach	Clontech	STO001	
4 -1	Claste -h	TTI 4 002	995 1171 61 219-220 273-276 312 315 330 596
thalamus	Clontech	THA002	963 996-1007 1059 1093 1160-1162
sh same	Clanatach	TUMOOL	8 120 151 208 221 316-317 353 639
thymus	Clonetech	THM001	750 867 874 878-881 927 963 1023
			1083 1094-1096 1124
thamasa	Clontech	THMc02	8 61 114 129 132 210 225 231 306
thymus	Cioniccii	1 LTMC05	317-319 336 340 359 380 398 446 448-
	1		463 512 519 545 554 587 598 698 724-
			725 789 812 836 868 873 927 947 952

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			976 1007 1042 1083 1085 1097-1116
			1122 1147 1177 1226-1229 1234 1311
L			1313
thyroid gland	Clontech	THR001	14 41 49 76 94 111 144 151 183 188
			210 217 222 253 264 271 277-286 294
			320-326 345-352 361 381-382 446 467
			483 514 534 549-550 564 578 602 649
	1		844 882-883 927 950 956 976 1008-
	1	1	1028 1076 1083 1117-1120 1142 1163-
<u></u>			1175 1230-1238 1308
trachea	Clontech	TRC001	223-225 238 287 353-354 514
			545 592 611 873 883-884 927
1			952 1029-1031 1042 1151-1152
			1170 1176-1177 1239
uterus	Clontech	UTR001	151 226 288-290 355 537 877
1	(885-886 976 1001 1032-1033
			1232

TABLE 2

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1	B02829	Homo sapiens	Human G protein coupled receptor hRUP5 protein SEQ ID NO:10.	460	100
2	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	111	51
3	R26173	Homo sapiens	Part of Major Yo paraneoplastic antigen (CDR62) encoded by clone pY2.	293	76
4	L29536	Homo sapiens	calcium channel L-type alpha 1 subunit	191	65
5	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	251	50
6	M11507	Homo sapiens	transferrin receptor	120	95
7	AF099100	Homo sapiens	WD-repeat protein 6	1941	93
8	Y92338	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-45.	245	82
9	G01343	Homo sapiens	Human secreted protein, SEQ ID NO: 5424.	226	91
10	AJ133798	Homo sapiens	copine VII protein	1127	68
11	G02449	Homo sapiens	Human secreted protein, SEQ ID NO: 6530.	584	99
12	X98330	Homo sapiens	ryanodine receptor 2	282	78
13	AL024498	Homo sapiens	dJ417M14.2 (novel serine/threonine-protein kinase (ortholog of mouse and rat MAK (male germ cell-associated kinase))	293	100
14	AF045577	Pan troglodytes	olfactory receptor OR93Ch	191	36
15	G03131	Homo sapiens	Human secreted protein, SEQ ID NO: 7212.	93	39
16	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein precursor	569	89
17	B08918	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:75.	99	44
18	Y36203	Homo sapiens	Human secreted protein #75.	165	75
19	U15647	Mus musculus	reverse transcriptase	106	40
20	G02701	Homo sapiens	Human secreted protein, SEQ ID NO: 6782.	544	100
21	Y35923	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 172.	1691	100
22	G04030	Homo sapiens	Human secreted protein, SEQ ID NO: 8111.	380	96
23	G02455	Homo sapiens	Human secreted protein, SEQ ID NO: 6536.	123	50
24	AF036329	Homo sapiens	gonadotropin-releasing hormone precursor, second form	284	90
25	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	96	32
26	S80119	Rattus sp.	reverse transcriptase homolog	100	34
27	U83303	Homo sapiens	line-1 reverse transcriptase	101	35
28	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	135	45

SEQ ID	Accession No.	Species	Description	Smith- Waterman Score	% Identity
NO:	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	83	42
29	G04087	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	116	72
30			Human secreted protein, SEQ ID NO: 7452.	96	67
31	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452. Human secreted protein, SEQ ID NO: 7305.	58	32
32	G03224	Homo sapiens		2457	98
33	Y66688	Homo sapiens	Membrane-bound protein PRO1152.		95
34	Y87071	Homo sapiens	Human secreted protein sequence SEQ ID NO:110.	348	
35	U15131	Homo sapiens	p126	182	48
36	Y73464	Homo sapiens	Human secreted protein clone yl4_1 protein sequence SEQ ID NO:150.	982	90
37	AL133215	Homo sapiens	bA108L7.6 (semaphorin 4G (sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain))	687	99
38	AC067969	amino acids 3338-4088	Homo sapiens ryanodine receptor 1 (skeletal)	386	66
39	AL031588	Homo sapiens	dJ1163J1.1 (mostly supported by GENSCAN, FGENES and GENEWISE)	493	76
40	G03628	Homo sapiens	Human secreted protein, SEQ ID NO: 7709.	110	51
41	AF132969	Homo sapiens	CGI-35 protein	228	68
42	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	220	88
42 43	X61048	Hydra sp.	mini-collagen	105	35
43	M76546	Helianthus	hydroxyproline-rich protein	110	31
		annuus Caenorhabditi	Rac-like GTPase	139	70
45	U82288	s elegans		118	58
46	G03477	Homo sapiens	Human secreted protein, SEQ ID NO: 7558.	113	63
47	AF090942	Homo sapiens	PRO0657	90	59
48	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	72	56
49	AJ005560	Mus musculus	SPR2B protein		
50	G02450	Homo sapiens	Human secreted protein, SEQ ID NO: 6531.	385	98
51	Y91649	Homo sapiens	Human secreted protein sequence encoded by gene 60 SEQ ID NO:322.	973	94
52	U93563	Homo sapiens	putative p150	105	38
53	Y55927	Homo sapiens	Human STLK2 protein.	699	85
54	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	145	56
55	AB008175	Mus musculus	hepatic nuclear factor 1-beta short form	356	74
56	M68941	Homo sapiens	protein-tyrosine phophatase	165	41
57	AL031600	Homo sapiens	c390E6.1 (chloride channel 7)	338	76
58	AF011417	Mus musculus	putative pheromone receptor	143	55
59	AF167320	Mus musculus	zinc finger protein ZFP113	558	68
60	U73036	Homo sapiens	interferon regultory factor 7	263	96
61	X07984	Mus musculus	protein-tyrosine kinase	297	69
62	Y29861	Homo sapiens	Human secreted protein clone cb98_4.	791	98
63	U35376	Homo sapiens	repressor transcriptional factor	485	65
64	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	785	74
65	G03883	Homo sapiens	Human secreted protein, SEQ ID NO: 7964.	88	95
66	AF177390	Manduca sexta	antennal specific membrane protein AMP	274	54
67	AB040800	Homo sapiens	SREB2	614	100
68	AF030027	Equine herpesvirus 4	24	213	26
69	G02965	Homo sapiens	Human secreted protein, SEQ ID NO: 7046.	261	95
70			Human oxidoreductase YTFO3.	1144	98
	W75770	Homo sapiens	KIAA0563 protein	239	76
71 72	AB011135 AB014885	Homo sapiens Halocynthia	HrPOPK-1	813	78
	AF045454	roretzi Cavia	phospholipase B	955	73
73	דכדכיי נה	porcellus	1		

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:				Score	Identity
	370000	musculus			ļ
75	Y00826	Rattus norvegicus	gp210 (AA 1-1886)	413	84
76	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	351	54
77	Y38422	Homo sapiens	Human secreted protein.	468	76
78	Y14596	Homo sapiens	Human T-type voltage-gated Ca channel alpha- 1-I (hCavT3).	1357	99
79	Y14591	Human papillomaviru s type 68	APM-1 protein	767	100
80	AL137802	Homo sapiens	dJ798A10.2 (KIAA0445 protein)	71	34
81	AP000383	Arabidopsis thaliana	protein arginine N-methyltransferase-like protein	359	65
82	L46815	Mus musculus	DNA binding protein Rc	895	75
83	G01600	Homo sapiens	Human secreted protein, SEQ ID NO: 5681.	315	96
84	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	538	71
85	AB029002	Homo sapiens	KIAA1079 protein	134	42
86	Y28678	Homo sapiens	Human cw272_7 secreted protein.	325	62
87	Y99368	Homo sapiens	Human PRO1326 (UNQ686) amino acid sequence SEQ ID NO:100.	156	48
88	AJ225124	Mus musculus	hyperpolarization-activated cation channel, HAC3 -	487	95
89	AF177203	Homo sapiens	cerebral cell adhesion molecule	290	56
90 91	Y28280 L39891	Homo sapiens	Human G-protein coupled receptor GRIR-2.	326	79
92	AF064876	Homo sapiens Homo sapiens	polycystic kidney disease-associated protein ion channel BCNG-1	1751 953	95
93	AF170723	Homo sapiens	protein kinase STK10	401	99
94 .	X13292	Trypanosoma brucei	GPI-phospholipase C (AA 1 - 358)	151	37
95	Y34127	Homo sapiens	Human potassium channel K+Hnov11.	661	99
96	X03638	Rattus norvegicus	sodium channel protein I (aa 1-2009)	1775	92
97	AF134213	Homo sapiens	ubiquitin-specific protease	1995	99
98	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	213	38
99	AF021935	Rattus norvegicus	mytonic dystrophy kinase-related Cdc42-binding kinase	675	48
100	AF279265	Homo sapiens	putative anion transporter 1	867	98
101	AC007878	Homo sapiens	match to nuclear protein, NP220; note: sequence difference at residue 58	160	60
102	U22829	Mus musculus	P2Y purinoceptor	264	42
103	Y45023	Homo sapiens	Human sensory transduction G-protein coupled receptor-B3.	516	99
104	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	787	98
105	Y87342	Homo sapiens	Human signal peptide containing protein HSPP- 119 SEQ ID NO:119.	343	57
106	AF169312	Homo sapiens	hepatic angiopoietin-related protein	212	67
107	AF116657	Homo sapiens	PRO1310	74	52
108	AE000401	Escherichia coli	sialic acid transporter	587	96
109	Y38395	Homo sapiens	Human secreted protein encoded by gene No. 10.	693	100
110	Y78801	Homo sapiens	Hydrophobic domain containing protein clone HP00631 amino acid sequence.	182	94
111	Z25535	Homo sapiens	nuclear pore complex protein hnup153	464	85
112	Y94939	Homo sapiens	Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84.	274	51
113	AF016365	Homo sapiens	hexokinase 1 isoform td	301	71
114	AC007956	Homo sapiens	unknown	520	75
115	M83738	Homo sapiens	protein-tyrosine phosphatase	251	92
	AL157952	Homo sapiens	dJ875K15.1.1 (ets homologous factor (ets- domain transcription factor ESE-3A, isoform 1))	484	91

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman Score	Identity
NO: 118	L41816	Homo sapiens	cam kinase I	407	62
119	AJ006710	Rattus	phosphatidylinositol 3-kinase	627	93
_		norvegicus		1646	94
120	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor, PDPr		
121	S39392	Homo sapiens	protein tyrosine phosphatase, PTPase {EC 3.1.3.48}	373	68
122	U60805	Homo sapiens	oncostatin-M specific receptor beta subunit	262	88
123	Y44403	Homo sapiens	Human truncated tankyrase-1.	111	35
124	Ú88167	Caenorhabditi s elegans	contains similarity to C2 domains	219	29
125	AF300648	Homo sapiens	guanine nucleotide binding protein beta subunit 4	693	90
126	AB021861	Mus musculus	apoptosis signal-regulating kinase 2	153	65
127	AF305210	Homo sapiens	concentrative Na+-nucleoside cotransporter hCNT3	807	97
128	M90360	Homo sapiens	protein kinase	220	73
129	D32202	Homo sapiens	alpha 1C adrenergic receptor isoform 2	574	86
130	AF208043	Homo sapiens	IFI16b	496	67
131	AF201734	Mus musculus	testis specific serine kinase-3	800	87
132	AF112886	Bos taurus	differentiation enhancing factor 1	159	74
133	AJ278314	Homo sapiens	phospholipase C-beta-1b	554	85
134	W74802	Homo sapiens	Human secreted protein encoded by gene 73 clone HSQEL25.	1157	87
135	AB020335	Homo sapiens	Pancreas-specific gene	668	96
136 137	W80408 AC002563	Homo sapiens Homo sapiens	A secreted protein encoded by clone dt674_2. putative RHO/RAC effector protein; 95%	866 5041	98
138	Y96736	Homo sapiens	similarity to P49205 (PID:g1345860) PRO3434, a novel secreted protein.	891	100
139	AB024034	Arabidopsis thaliana	DNA-damage inducible protein DDI1-like	147	55
140	W97809	Homo sapiens	Human GTPase regulator GRAF.	248	56
141	Y51557	Homo sapiens	Human PLA2 protein.	125	46
142	AF090113	Rattus norvegicus	AMPA receptor binding protein	623	93
143	W26642	Homo sapiens	Human RECK cancer-inhibiting protein.	641	82
144	U87306	Rattus norvegicus	transmembrane receptor UNC5H2	578	84
145	AF264014	Homo sapiens	scavenger receptor cysteine-rich type I protein M160 precursor	727	92
146	W63683	Homo sapiens	Human secreted protein 3.	140	40
147 148	M96264 D64014	Homo sapiens Escherichia	galactose-1-phosphate uridyl transferase HrsA	513 818	81 90
149	M83316	Escherichia coli	pppGpp phosphohydrolase	915	95
150	AL163279	Homo sapiens	homolog to cAMP response element binding and beta transducin family proteins	1261	99
151	AF179867	Homo sapiens	STE20-like kinase	940	99
152	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain ligand (clone 3TW).	392	61
153	AF151859	Homo sapiens	CGI-101 protein	370	92
154	X66957	Homo sapiens	hexokinase type I	489	81
155	Y16355	Homo sapiens	alternatively spliced form	432	92
156	G00857	Homo sapiens	Human secreted protein, SEQ ID NO: 4938.	349	78
157	AF159455	Mus musculus	zinc finger protein	352	74
158	L76191	Homo sapiens	interleukin-1 receptor-associated kinase	537	76
159	AP001743	Homo sapiens	putative gene, ankirin like, possible dual specifity Ser/Thr/Tyr kinase domain	670	98
160	AJ250425	Rattus norvegicus	Collybistin I	556	74
161	G02885	Homo sapiens	Human secreted protein, SEQ ID NO: 6966.	370	100

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	open.co		Waterman	Identity
NO:	1	1		Score	luchary
162	Z22968	Homo sapiens	M130 antigen	610	100
163	AF181121	Homo sapiens	ATP-dependent Ca2+ pump PMR1	336	92
164	AF055636	Homo sapiens	leucine-rich glioma-inactivated protein precursor	455	94
165	AF160798	Rattus	calcium transporter CaT1	700	96
		norvegicus	·		
166	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	327	45
167	Y48607	Homo sapiens	Human breast tumour-associated protein 68.	1072	99
168	AB020741	Mus musculus	NIK-related kinase	197	43
169	AF252293	Homo sapiens	PAR3	596	44
170	U59429	Cricetinae gen. sp.	diacylglycerol kinase eta	481	82
171	AF035268	Homo sapiens	phosphatidylserine-specific phospholipase A1	386	42
172	AF127085	Mus	semaphorin cytoplasmic domain-associated	507	82
173	Y27918	musculus Homo sapiens	Protein 3B Human secreted protein encoded by gene No.	653	99
151	1000000	 	123.		\ <u></u>
174	G02979	Homo sapiens	Human secreted protein, SEQ ID NO: 7060.	538	97
175	U36488	Mus musculus	embryonic stem cell phosphatase	168	55
176	W95629	Homo sapiens	Homo sapiens secreted protein gene clone gm196_4.	1022	100
177	AF289023	Homo sapiens	formiminotransferase cyclodeaminase form D	255	93
178	X04936	Homo sapiens	T-cell receptor alpha-chain (413 is 2nd base in codon)	710	99
179	AF127481	Homo sapiens	non-ocogenic Rho GTPase-specific GTP exchange factor	175	80
180	G00978	Homo sapiens	Human secreted protein, SEQ ID NO: 5059.	517	94
181	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	671	96
182	AF110640	Homo sapiens	orphan seven-transmembrane receptor	862	100
183	AB020854	Bos taurus	orphan transporter short splicing variant	766	84
184	AF169691	Homo sapiens	cadherin-like protein VR8	375	38
185	AF126372	Homo sapiens	thyrotropin-releasing hormone degrading ectoenzyme	985	99
186	L20966	Homo sapiens	phosphodiesterase	541	76
187	G02920	Homo sapiens	Human secreted protein, SEQ ID NO: 7001.	254	93
188	Y94918	Homo sapiens	Human secreted protein clone dd504_18 protein	301	98
189	Y66713	Homo sapiens	sequence SEQ ID NO:42. Membrane-bound protein PRO1309.	694	100
190	G03244		Human secreted protein, SEQ ID NO: 7325.	331	73
191	U36771	Rattus norvegicus	sn-glycerol 3-phosphate acyltransferase	707	92
192	R05935	Homo sapiens	Secreted GPIIb subunit of multiple subunit polypeptide (MSP)GPIIb-IIIa.	157	72
193	M92084	Theileria parva	casein kinase II alpha subunit	364	50
194	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	448	90
195	W95631	Homo sapiens	Homo sapiens secreted protein gene clone	382	49
196	AF255614	Rattus norvegicus	hj968_2. scaffolding protein SLIPR	680	99
197	AC021640	Arabidopsis	putative phosphatidate phosphohydrolase	300	41
198	AF073967	thaliana Mus musculus	olfactory receptor	316	43
		domesticus			
199	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	617	98
200	AF117948	Homo sapiens	pancreas-enriched phospholipase C	625	89
201	AF128625	Homo sapiens	CDC42-binding protein kinase beta	636	94
202	AF117946	Homo sapiens	Link guanine nucleotide exchange factor II	1303	100
203	Y53021	Homo sapiens	Human secreted protein clone qc646_1 protein sequence SEQ ID NO:48.	701	99
203			L SCHULIG SELVIDINI PAN		
204	AF227968	Homo sapiens	SH2-B beta signaling protein	182	79

SEQ	Accession	Species	Description	Smith-	1%
ID T	No.	Openies	, Description	Waterman	Identity
NO:	110.	1		Score	12011119
140.	 	 	{ovarian cancer critical region of deletion}	-	
206	U18315	Sus scrofa	parathyroid receptor	122	60
207	AF255342	Homo sapiens	putative pheromone receptor V1RL1 long form	170	96
	1			715	94
208	S52051	Rattus sp.	neurotransmitter transporter		
209	W63683	Homo sapiens	Human secreted protein 3.	840	99
210	D79992	Homo sapiens	similar to Drosophila photoreceptor cell-specific	541	82
	<u> </u>		protein, calphotin.		
211	AF117948	Homo sapiens	pancreas-enriched phospholipase C	1348	99
212	U81035	Rattus	ankyrin binding cell adhesion molecule	471	69
		norvegicus	neurofascin		
213	AF154846	Homo sapiens	zinc finger protein	798	56
214	AF102777	Mus	FYVE finger-containing phosphoinositide kinase	933	93
	Í	musculus			
215	AL163303	Homo sapiens	putative gene containing transmembrane domain	523	89
216	U26595	Rattus	prostaglandin F2a receptor regulatory protein	563	78
	1	norvegicus	precursor		1
217	G04095	Homo sapiens	Human secreted protein, SEQ ID NO: 8176.	644	98
218	X75756	Homo sapiens	protein kinase C mu	314	81
219	Y66723	Homo sapiens	Membrane-bound protein PRO1100.	770	98
220	D88577	Mus	Kupffer cell receptor	567	40
44 U	D003//	musculus	Kupitei celi iccepioi	1 307) **
221	AF258465		OTRPC4	853	100
221		Homo sapiens		636	96
222	AF021935	Rattus	mytonic dystrophy kinase-related Cdc42-binding	636	96
	<u> </u>	norvegicus	kinase	-	100
223	AL136527	Homo sapiens	bA215B13.1 (A kinase (PRKA) anchor protein	693	100
		<u> </u>	11)	(00	100
224	AB032417	Homo sapiens	WNT receptor Frizzled-4	690	99
225	AF030430	Mus	semaphorin VIa	703	68
		musculus			
226	AE000218	Escherichia	putative dihydroxyacetone kinase (EC 2.7.1.2)	297	39
	<u> </u>	coli		<u>L</u>	<u> </u>
227	AF302150	Homo sapiens	phosphoinositol 3-phosphate-binding protein-2	2080	100
228	AB024573	Mus	GTP-binding like protein 2	265	88
		musculus			
229	AF122924	Xenopus	Wnt inhibitory factor-1	316	40
		laevis			1
230	G03205	Homo sapiens	Human secreted protein, SEQ ID NO: 7286.	229	100
231	X98260	Homo sapiens	M-phase phosphoprotein 11	265	92
232	R92754	Homo sapiens	Human growth differentiation factor-12.	682	95
233	R75111	Homo sapiens	Glycosyl-phosphatidylinositol-specific	290	100
200	1,7,7,7,7	120me suprem	phospholipase-D.		
234	W69431	Homo sapiens	Human secreted protein cw1233_3.	235	97
	Y08686		serine palmitoyltransferase, subunit II	859	81
235	1	Homo sapiens Homo sapiens		117	37
236	AF118275		atrophin-related protein ARP		62
237	X81466	Mus	Embryo Brain Kinase	460	02
	1	musculus			-
238	U64857	Caenorhabditi	similar to the BPTI/Kunitz family of inhibitors;	284	33
		s elegans	most similar to tissue factor pathway inhibitor	İ	
		<u> </u>	precursor (TFPI)	L	<u> </u>
239	AJ250840	Mus	serine/threonine protein kinase	739	63
	<u> </u>	musculus			
240	AJ223472	Mus	transcription elongation factor TFIIS.h	222	38
		musculus			
241	Y94906	Homo sapiens	Human secreted protein clone rb649_3 protein	353	52
		1	sequence SEQ ID NO:18.		
242	AF169301	Homo sapiens	Na+/sulfate cotransporter SUT-1	591	99
243	L22022	Rattus	orphan transporter v7-3	667	93
		norvegicus		}	1
244	AF016191	Rattus	potassium channel	1043	98
277	71 010171	norvegicus	potassium enamior	1045	1 20
	AF097366	Homo sapiens	cone sodium-calcium potassium exchanger	645	98
7/5	1 WEAN 1200		Human secreted protein clone pp325_9.	497	98
245	3720070		i muman secreteo protein Clone DD1/3 Y	1 44/	ו אל ו
246	Y29868	Homo sapiens			02
246 247	AF180475	Homo sapiens	Not4-Np	188	83
246					83 99 31

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
NO.	 	sexta	protein SCLP	Score	
250	AF192756	Kaposi's	Orf73	134	34
250	AL 192750	sarcoma- associated	011/3	, 134	34
		herpesvirus			}
251	AB022694	Homo sapiens	MOK protein kinase	209	83
252	W55045	Homo sapiens	Neural adhesion molecule (ethb0018f2 product).	469	100
253	L46815	Mus musculus	DNA binding protein Rc	251	67
254	W68505	Homo sapiens	Human acid sensing ionic channel.	173	82
255	AF070066	Mus musculus	Citron-K kinase	1201	98
256	G02491	Homo sapiens	Human secreted protein, SEQ ID NO: 6572.	460	100
257	Z12841	Oryctolagus cuniculus	Phospholipase	368	80
258	Y95436	Homo sapiens	Human calcium channel SOC-3/CRAC-2.	1857	99
259	AJ222968	Mus musculus	L-periaxin	430	72
260	AJ250839	Homo sapiens	serine/threonine protein kinase	861	100
261	AJ249977	Homo sapiens	AMP-activated protein kinase gamma 3 subunit	758	98
262	AF141386	Rattus norvegicus	SLIT-2	198	40
263	AF022859	Homo sapiens	neuropilin-2(a0)	335	62
264	AF160477	Homo sapiens	Ig superfamily receptor LNIR precursor	387	91
265	Y44662	Homo sapiens	Human 14273 G-protein coupled receptor (GPCR).	636	99
266	U27269	Mus musculus	sodium glucose cotransporter	204	56
267	AF124491	Homo sapiens	ARF GTPase-activating protein GIT2	159	75
268	AF127389	Rattus norvegicus	putative taste receptor TR1	209	39
269	X98296	Homo sapiens	ubiquitin hydrolase	215	95
270	X78482	Streptococcus pyogenes	Fc-gamma receptor	129	26
271	AB009883	Nicotiana tabacum	KED	109	26
272	AF137367	Mus musculus	VPS10 domain receptor protein SORCS	899	97
273	L34938	Rattus norvegicus	ionotropic glutamate receptor	460	86
274	AL022724	Homo sapiens	dJ413H6.1.1 (hamster Androgen-dependent Expressed Protein LIKE PUTATIVE protein) (isoform 1)	188	74
275	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	173	94
276	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	148	56
277	L40380	Homo sapiens	thyroid receptor interactor	430	61
278	AB046851	Homo sapiens	KIAA1631 protein	283	96
279	AC008075	Arabidopsis thaliana	Contains PF 00069 Eukaryotic protein kinase domain.	157	43
280	M83738	Homo sapiens	protein-tyrosine phosphatase	181	73
281	AK024397	Homo sapiens	unnamed protein product	439	91
282	AF141326	Homo sapiens	RNA helicase HDB/DICE1	497	84
283	AF156530	Mus musculus	ETS-domain transcriptional repressor PE1	605	76
284	Y29336	Homo sapiens	Human secreted protein clone cs756_2 alternate reading frame protein.	647	100
285	Y73402	Homo sapiens	Human secreted protein clone yc25_1 protein sequence SEQ ID NO:26.	300	90
286	AF016411	Homo sapiens	KCNA3.1B	137	100
287	W89253	Homo sapiens	Human ALP.	688	97
288	AF112886	Bos taurus	differentiation enhancing factor 1	750	96
289	AF113131	Homo sapiens	host cell factor homolog LCP	367	44
290	U52111	Homo sapiens	plexin-related protein	698	100
291	AF026504	Rattus	SPA-1 like protein p1294	603	89

SEQ	Accession	Species	Description	Smith-	1 %
ID	No.	Species	Description	Waterman	Identity
	No.	İ		Score	Identity
NO:	 	norvegicus		30016	
-	AF102854			124	53
292	AF 102854	Rattus	membrane-associated guanylate kinase-	124	23
	1,100000	norvegicus	interacting protein 2 Maguin-2		1 20
293	X99211	Drosophila	ubiquitin-specific protease	143	38
	<u> </u>	melanogaster			
294	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein	185	94
		<u> </u>	sequence SEQ ID NO:92.		<u> </u>
295	Y94890	Homo sapiens	Human protein clone HP02798.	108	59
296	AF019767	Homo sapiens	zinc finger protein	154	96
297	Y28568	Homo sapiens	Secreted peptide clone bd577_1.	568	84
298	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein	182	97
		<u> </u>	sequence SEQ ID NO:92.		
299	B08906	Homo sapiens	Human secreted protein sequence encoded by	605	69
			gene 16 SEQ ID NO:63.		
300	R58890	Homo sapiens	Human-32 cadherin-related molecule.	212	97
301	AF022859	Homo sapiens	neuropilin-2(a0)	277	100
302	Y71124	Homo sapiens	Human mitogenic regulator duox2.	716	97
303	Y44297	Homo sapiens	Human receptor tyrosine kinase.	228	97
304	D32050	Homo sapiens	alanyl-tRNA synthetase	192	80
305	U43586	Homo sapiens	protein kinase related to Raf protein kinases;	428	72
	1		Method: conceptual translation supplied by		
		j	author		1
306	R54872	Homo sapiens	Human H13 viral receptor mutant 4.	280	95
307	D78572	Mus	membrane glycoprotein	199	41
50,	7,00.2	musculus	mand gry voprotean		''
308	AF255614	Rattus	scaffolding protein SLIPR	639	88
		norvegicus	5		**
309	S79463	Mus sp.	semaphorin homolog-M-Sema F	162	89
310	AF178941	Homo sapiens	ATP-binding cassette sub-family A member 2	736	100
311	U03413	Dictyostelium	calcium binding protein	151	36
311	003413	discoideum	calcium omanig protein	131	30
312	¥87347	Homo sapiens	Human signal peptide containing protein HSPP-	744	100
312	10/54/	110mo sapiciis	124 SEO ID NO:124.	/ ***	100
313	Z97055	Homo sapiens	dJ388M5.4 (putative GS2 like protein)	789	99
314	AC004010	Homo sapiens	similar to Leucine-rich transmembrane proteins;	197	38
314	ACCOTOLO	Tromo sapicus	44% similarity to U42767 (PID:g1736918)	.,,	1 30
315	AL021392	Homo sapiens	dJ439F8.2 (supported by GENSCAN and	278	38
313	7110021352	110mo sapiciis	GENEWISE)	270	1 30
316	U70209	Mus	polycystic kidney disease I protein	165	38
310	070205	musculus	polycystic kidney disease i protein	103	36
317	AF109643	Rattus	coxsackie-adenovirus-receptor homolog	223	38
317	74 102043	norvegicus	coxsackie-adenovirus-receptor nomolog	223	30
318	AF104923	Homo sapiens	putative transcription factor	138	84
319	AF100287	Trypanosoma	activated protein kinase C receptor homolog	141	38
319	AI 100287	vivax	activated protetti kinase C receptor nomolog	141	36
320	G00588	Homo sapiens	Human secreted protein, SEQ ID NO: 4669.	125	51
321	Y21591	Homo sapiens	Human secreted protein (clone CC332-33).	459	97
322	D26070		human type 1 inositol 1,4,5-trisphosphate	232	97
344	1020070	Homo sapiens		232	} "
222	V27019	IIIama andina	receptor	306	00
323	Y27918	Homo sapiens	Human secreted protein encoded by gene No.	306	88
224	AE010144	TTages	123.	200	170
324	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	209	70
325	M19650	Homo sapiens	2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC	214	97
206	1 111005.03	 	3.1.4.37)	140	+
326	W80396	Homo sapiens	A secreted protein encoded by clone bp646_10.	140	70
327	X75756	Homo sapiens	protein kinase C mu	540	78
328	G02292	Homo sapiens	Human secreted protein, SEQ ID NO: 6373.	721	99
329	AF168990	Homo sapiens	putative GTP-binding protein	877	99
330	S67984	Homo sapiens	anti-HIV gp120 antibody heavy chain variable	581	80
			region		
331	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	2823	98
332	Y87330	Homo sapiens	Human signal peptide containing protein HSPP-	1127	100
	<u> </u>	<u> </u>	107 SEQ ID NO:107.		<u></u>
333	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	320	98
334	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95%	327	93

SEQ	Accession No.	Species	Description	Smith-	%
ID NO:	No.			Waterman Score	Identity
10.	 	-	similarity to P49205 (PID:g1345860)	Score	+
335	Y87347	Homo sapiens	Human signal peptide containing protein HSPP-	1111	67
		<u> </u>	124 SEQ ID NO:124.		<u> </u>
336	AF006466	Mus musculus	lymphocyte specific formin related protein	193	75
337	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	632	97
338	Y13443	Homo sapiens	Amino acid sequence of hSlo3-2.	516	100
339	Y07637	Homo sapiens	putative GABA-gated chloride channel	189	100
340	Y05734	Homo sapiens	Human Grb7 effector 2.2412 protein.	2156	99
341	AE000497	Escherichia coli	L-idonate transcriptional regulator	928	98
342	D90855	Escherichia coli	glycerol-3-phosphate dehydrogenase (EC 1.1.99.5) chain A, anaerobic	769	99
343	D85613	Escherichia coli	membrane component	399	100
344	M93239	Escherichia coli	transmembrane protein	232	100
345	M60177	Escherichia coli	enterobactin	759	99
346	D90699	Escherichia coli	Sensor protein copS (EC 2.7.3).	638	97
347	D90843	Escherichia coli	CapB protein.	552	100
348	M13422	Escherichia coli	49 kd protein	1193	96
349	L10328	Escherichia coli	similar to drug resistance translocases	340	90
350	X69942	Mus musculus	enhancer-trap-locus-1	560	82
351	AF239613	Homo sapiens	apamin-sensitive small-conductance Ca2+-	463	80
	1.2.250.3	Aromo suprems	activated potassium channel	.05	"
352	D90777	Escherichia coli	3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157) (b- hydroxybutyryl-CoA dehydrogenase) (BhbD).	577	100
353	D90863	Escherichia coli	similar to	311	98
354	Y52386	Homo sapiens	Human transmembrane protein HP02000.	133	58
355	Y31645	Homo sapiens	Human transport-associated protein-7 (TRANP-7).	482	55
356	Y58637	Homo sapiens	Protein regulating gene expression PRGE-30.	119	51
357	AF119226	Homo sapiens	dual-specificity tyrosine phosphatase YVH1	1788	100
358	Y87219	Homo sapiens	Human secreted protein sequence SEQ ID NO:258.	165	100
359	J00132	Homo sapiens	beta-fibrinogen	233	93
360	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	128	70
361	R28916	Homo sapiens	Type III procollagen (prior art).	108	40
362	U16655	Rattus norvegicus	phospholipase C delta-4	649	65
363	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	95	42
364	U47276	Gallus gallus	chicken brain factor-2	104	34
365	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	183	65
366	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	118	46
367	X98258	Homo sapiens	M-phase phosphoprotein 9	564	75
368	AL021366	Homo sapiens	cICK0721Q.3 (Kinesin related protein)	3387	99
369	U70932	Peromyscus leucopus	reverse transcriptase	92	5,9
370	X86400	Homo sapiens	gamma subunit of sodium potassium ATPase like	242	73
371	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	165	56
372	U49974	Homo sapiens	mariner transposase	257	55
373	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	21193	99
374	AF234765	Rattus norvegicus	serine-arginine-rich splicing regulatory protein SRRP86	1182	78
375	U49974	Homo sapiens	mariner transposase	172	55

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	1	1 '		Score	
376	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	221	67
377	G00669	Homo sapiens	Human secreted protein, SEQ ID NO: 4750.	600	100
378	X52574	Mus	GTP binding protein	1456	91
	1122011	musculus	·		
379	R69095	Homo sapiens	Anti-HIV Fab tat31 light chain.	68	37
380	J04974	Homo sapiens	alpha-2 type XI collagen	125	37
381	AB002405	Homo sapiens	LAK-4p	530	43
382	U64830	Dictyostelium	protein tyrosine kinase	115	44
		discoideum	provins gradual scales	1	'
383	G02916	Homo sapiens	Human secreted protein, SEQ ID NO: 6997.	618	98
384	G01194	Homo sapiens	Human secreted protein, SEQ ID NO: 5275.	617	93
385	AJ245822	Homo sapiens	type I transmembrane receptor	4560	100
386	D86974	Homo sapiens	KIAA0220	2148	98
387	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	142	50
388	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	99	59
389	M12140	Homo sapiens	envelope protein	197	51
390	AJ293309	Homo sapiens	NHP2 protein	461	77
391	Y42751	Homo sapiens	Human calcium binding protein 2 (CaBP-2).	181	94
392	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	241	66
393	Y14442	Homo sapiens	olfactory receptor protein	339	54
394	W85607	Homo sapiens	Secreted protein clone da228_6.	957	100
395	Y76332	Homo sapiens	Fragment of human secreted protein encoded by	171	34
	170552	TIOINO Sapions	gene 38.	1/1] 34
396	G03930	Homo sapiens	Human secreted protein, SEQ ID NO: 8011.	250	100
397	AB032904	Hylobates	dopamine receptor D4	105	35
371	130032304	syndactylus	dopainine receptor D4	103	33
398	AJ007798	Homo sapiens	stromal antigen 3, (STAG3)	861	.85
399	Y91405	Homo sapiens	Human secreted protein sequence encoded by	1047	92
377	151-105	Tiomo sapions	gene 2 SEQ ID NO:126.	1047	~~
400	Y29861	Homo sapiens	Human secreted protein clone cb98 4.	162	37
401	D87002	Homo sapiens	similar to rat integral membrane glycoprotein;	527	78
401	207002	Tiomo supiens	accession number Z21513.	1 321	1 10
402	AF100754	Homo sapiens	ancient ubiquitous protein AUP1 isoform	853	95
403	X74904	Gallus gallus	alpha-2-macroglobulin receptor	258	60
404	AF075462	Mus	ADP-ribosylation factor-directed GTPase	545	89
	12070102	musculus	activating protein isoform b	1	"
405	X92887	Human	pol/env	162	30
		endogenous	F		1
	1	retrovirus K		{	ļ
406	Y30162	Homo sapiens	Human dorsal root receptor 4 hDRR4.	325	72
407	AK022626	Homo sapiens	unnamed protein product	2833	99
408	L13802	Homo sapiens	ribosmal protein small subunit	264	92
409	Y91600	Homo sapiens	Human secreted protein sequence encoded by	1788	89
	}	•	gene 9 SEQ ID NO:273.		,
410	W88745	Homo sapiens	Secreted protein encoded by gene 30 clone	2004	99
ı	1		HTSEV09.		
411	AB043953	Mus	Chat-H	2628	82
		musculus			1
				1	+==
412	Y86233	Homo sapiens	Human secreted protein HNTMX29, SEQ ID	1014	92
412	Y86233	Homo sapiens	Human secreted protein HNTMX29, SEQ ID NO:148.	1014	92
		Homo sapiens	NO:148.	265	
412	Y86233 U10542			l	92
		Pan	NO:148. MHC class I A	l	
413	U10542	Pan troglodytes	NO:148. MHC class I A NY-REN-7 antigen	265	71
413	U10542 AF155097 G03203	Pan troglodytes Homo sapiens Homo sapiens	NO:148. MHC class I A NY-REN-7 antigen Human secreted protein, SEQ ID NO: 7284.	265 850	71
413 414 415 416	U10542 AF155097 G03203 Y57911	Pan troglodytes Homo sapiens Homo sapiens Homo sapiens	NO:148. MHC class I A NY-REN-7 antigen Human secreted protein, SEQ ID NO: 7284. Human transmembrane protein HTMPN-35.	265 850 88 266	71 95 48 89
413 414 415 416 417	U10542 AF155097 G03203 Y57911 W27651	Pan troglodytes Homo sapiens Homo sapiens Homo sapiens	NO:148. MHC class I A NY-REN-7 antigen Human secreted protein, SEQ ID NO: 7284. Human transmembrane protein HTMPN-35. Secreted protein AT205.	265 850 88 266 481	71 95 48 89 60
413 414 415 416 417 418	U10542 AF155097 G03203 Y57911 W27651 Y76884	Pan troglodytes Homo sapiens Homo sapiens Homo sapiens Homo sapiens	NO:148. MHC class I A NY-REN-7 antigen Human secreted protein, SEQ ID NO: 7284. Human transmembrane protein HTMPN-35. Secreted protein AT205. Retinoblastoma binding protein-7 sequence.	265 850 88 266 481 3077	71 95 48 89 60 87
413 414 415 416 417	U10542 AF155097 G03203 Y57911 W27651	Pan troglodytes Homo sapiens Homo sapiens Homo sapiens Homo sapiens Notothenia	NO:148. MHC class I A NY-REN-7 antigen Human secreted protein, SEQ ID NO: 7284. Human transmembrane protein HTMPN-35. Secreted protein AT205.	265 850 88 266 481	71 95 48 89 60
413 414 415 416 417 418 419	U10542 AF155097 G03203 Y57911 W27651 Y76884 AF255559	Pan troglodytes Homo sapiens Homo sapiens Homo sapiens Homo sapiens Notothenia coriiceps	NO:148. MHC class I A NY-REN-7 antigen Human secreted protein, SEQ ID NO: 7284. Human transmembrane protein HTMPN-35. Secreted protein AT205. Retinoblastoma binding protein-7 sequence. alpha tubulin	265 850 88 266 481 3077 289	71 95 48 89 60 87 68
413 414 415 416 417 418 419	U10542 AF155097 G03203 Y57911 W27651 Y76884 AF255559 G01984	Pan troglodytes Homo sapiens Homo sapiens Homo sapiens Homo sapiens Notothenia coriiceps Homo sapiens	NO:148. MHC class I A NY-REN-7 antigen Human secreted protein, SEQ ID NO: 7284. Human transmembrane protein HTMPN-35. Secreted protein AT205. Retinoblastoma binding protein-7 sequence. alpha tubulin Human secreted protein, SEQ ID NO: 6065.	265 850 88 266 481 3077 289	71 95 48 89 60 87 68
413 414 415 416 417 418 419	U10542 AF155097 G03203 Y57911 W27651 Y76884 AF255559	Pan troglodytes Homo sapiens Homo sapiens Homo sapiens Homo sapiens Notothenia coriiceps	NO:148. MHC class I A NY-REN-7 antigen Human secreted protein, SEQ ID NO: 7284. Human transmembrane protein HTMPN-35. Secreted protein AT205. Retinoblastoma binding protein-7 sequence. alpha tubulin Human secreted protein, SEQ ID NO: 6065. dJ309K20.2 (acrosomal protein ACR55 (similar	265 850 88 266 481 3077 289	71 95 48 89 60 87 68
413 414 415 416 417 418 419	U10542 AF155097 G03203 Y57911 W27651 Y76884 AF255559 G01984	Pan troglodytes Homo sapiens Homo sapiens Homo sapiens Homo sapiens Notothenia coriiceps Homo sapiens	NO:148. MHC class I A NY-REN-7 antigen Human secreted protein, SEQ ID NO: 7284. Human transmembrane protein HTMPN-35. Secreted protein AT205. Retinoblastoma binding protein-7 sequence. alpha tubulin Human secreted protein, SEQ ID NO: 6065.	265 850 88 266 481 3077 289	71 95 48 89 60 87 68

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	1	 		Score	<u> </u>
423	AF231705	Homo sapiens	Alu co-repressor 1	1090	100
424	AF234887	Homo sapiens	FLAMINGO 1	6268	97
425	Y35942	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 191.	1961	99
426	AB009288	Homo sapiens	N-copine	635	98
427	L12392	Homo sapiens	Huntington's Disease protein	16080	99
428	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	768	98
429	AJ293573	Homo sapiens	zinc finger protein Cezanne	542	87
430	Y84441	Homo sapiens	Amino acid sequence of a human RNA-associated protein.	2074	100
431	G02850	Homo sapiens	Human secreted protein, SEQ ID NO: 6931.	723	95
432	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	73	42
433	AF159296	Lycopersicon esculentum	extensin-like protein	613	48
434	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	135	44
435	X73874	Homo sapiens	phosphorylase kinase	3442	97
436	AF161426	Homo sapiens	HSPC308	268	74
437	Y30812	Homo sapiens	Human secreted protein encoded from gene 2.	1055	52
438	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	168	56
439	X14766	Homo sapiens	GABA-A receptor alpha 1 subunit	2294	96
440	X02344	Homo sapiens	beta-tubulin	311	95
441	AF168418	Homo sapiens	activating signal cointegrator 1	1882	100
442	L11672	Homo sapiens	zinc finger protein	795	54
443	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	93	26
444	A52140	unidentified	HUMAN NDR	2451	100
445	X98330	Homo sapiens	ryanodine receptor 2	9356	99
446	AF116712	Homo sapiens	PRO2738	227	49
447	AF245447	Homo sapiens	sphingosine kinase type 2 isoform	576	99
448	AF133086	Homo sapiens	membrane-type serine protease 1	2630	94
449	U87305	Rattus norvegicus	transmembrane receptor UNC5H1	817	93
450	AF081249	Homo sapiens	JAW1-related protein MRVI1A long isoform	4568	99
451	AC005498	Homo sapiens	R31665_1	316	62
452	M60235	Homo sapiens	granule membrane protein-140	464	73
453	AB036706	Homo sapiens	intelectin	730	88
454	G00918	Homo sapiens	Human secreted protein, SEQ ID NO: 4999.	263	81
455	Y22634	Homo sapiens	Human cytokine inducible regulatory protein-1 (CIRP-1).	192	67
456	Y36705	Homo sapiens	Fragment of human secreted protein encoded by gene 62.	106	40
457	N91325	Homo sapiens	DNA encoding human growth hormone receptor.	3282	96
458	M19155	Plasmodium	S-antigen precursor	110	36
		falciparum			
459	Y13377	Homo sapiens	Amino acid sequence of protein PRO257.	509	98
460	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	149	43
461	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	184	54
462	Y53005	Homo sapiens	Human secreted protein clone pm749_8 protein sequence SEQ ID NO:16.	135	47
463	X84960	Triticum aestivum	low molecular weight glutenin	109	33
464	W19919	Homo sapiens	Human Ksr-1 (kinase suppressor of Ras).	1781	85
465	AF189764	Mus musculus	alpha/beta hydrolase-1	502	59
466	U93569	Homo sapiens	p40	101	30
467	Y41528	Homo sapiens	Fragment of human secreted protein encoded by gene 77.	1172	99
468	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	149	52
469	AJ000008	Homo sapiens	PI3-kinase	5832	97
470	X70922	Mus musculus	neurotoxin homologue	118	47
471	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	198	75
472	Y36705	Homo sapiens	Fragment of human secreted protein encoded by	72	57

SEQ	Accession	Species	Description	Smith-	%
ID `	No.			Waterman	Identity
NO:				Score	1
	1		gene 62.		1
473	G02313	Homo sapiens	Human secreted protein, SEQ ID NO: 6394.	328	100
474	Y07007	Homo sapiens	Breast cancer associated antigen precursor	1013	97
	1		sequence.	1	
475	W93254	Homo sapiens	Human ESRP1 protein.	943	80
476	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	236	65
477	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	202	60
	ł		clone HTDAD22.	1	i
478	G01870	Homo sapiens	Human secreted protein, SEQ ID NO: 5951.	267	100
479	AF102777	Mus	FYVE finger-containing phosphoinositide kinase	3427	92
		musculus			1
480	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	123	53
481	W87701	Homo sapiens	A human membrane fusion protein designated	221	77
			SYTAX1.	l	
482	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	131	39
483	AF210651	Homo sapiens	NAG18	124	59
484	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	343	50
485	G00637	Homo sapiens	Human secreted protein, SEQ ID NO: 4718.	129	70
486	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	149	73
	1		3		1
487	Y76167	Homo sapiens	Human secreted protein encoded by gene 44.	627	100
488	AJ275213	Homo sapiens	stabilin-1	1244	91
489	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	313	65
490	L12392	Homo sapiens	Huntington's Disease protein	16081	100
491	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	197	66
492	J03799	Homo sapiens	laminin-binding protein	228	70
493	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	128	41
	j		3 .		
494	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	197	67
			clone HTDAD22.	!	
495	AC005175	Homo sapiens	R31449_3	889	94
496	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	229	61
497	AB030237	Canis	D4 dopamine receptor	90	48
		familiaris			
498	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	228	65
499	U70935	Peromyscus	reverse transcriptase	213	52
		maniculatus		i	1
500	U48508	Homo sapiens	skeletal muscle ryanodine receptor	26406	99
501	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	105	58
502	AF119851	Homo sapiens	PRO1722	156	62
503	AF113685	Homo sapiens	PRO0974	116	50
504	U79458	Homo sapiens	WW domain binding protein-2	322	59
505	W29651	Homo sapiens	Human secreted protein CD124_3.	608	55
506	W85459	Homo sapiens	Secreted protein encoded by clone dh1135_9.	986	70
507	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID	115	33
			NO:180.	<u></u>	
508	AL160175	Homo sapiens	bA243J16.3 (similar to MYLK (myosin, light	184	92
			polypeptide kinase))	1	
509	U43360	Peromyscus	reverse transcriptase	97	62
	<u>L</u>	maniculatus		L	
510	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	117	63
511	W79092	Homo sapiens	Human secreted protein dn740_3.	1058	100
512	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	205	64
513	AJ133439	Homo sapiens	GRIP1 protein	2151	100
514	AE003456	Drosophila	CG6393 gene product	259	42 .
		melanogaster	<u> </u>	<u> </u>	1
515	Z17206	Xenopus	p46XIEg22	128	40
	İ	laevis	•		
516	AF104413	Homo sapiens	large tumor suppressor 1	1766	94
517	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	92	40
518	AF151083	Homo sapiens	HSPC249	444	98
		Homo sapiens	cytochrome c-like polypeptide	318	50
519	S80864	HOIMO SAPIGIIS	oytochlothe o-like polypopilas	1 2 1 0	
	X92485	Plasmodium	pval	170	61

Document September Septe	SEQ	Accession	Species	Description	Smith-	1%
No.	-		opecies	Description		
		110.				identity
		G03790	Homo caniene	Human secreted protein CEO ID NO. 2021		+
Supplementary Supplementar						
W88627		<u> </u>				
HPMBQ32.				Secreted protein, SEQ ID NO: 6/35.		
Human secreted protein, SEQ ID NO: 6788. 154 57 527 560707 Homo sapiens Human secreted protein, SEQ ID NO: 6788. 1112 86 529 564663 Homo sapiens C8 Human secreted protein, SEQ ID NO: 8144. 84 4111 55 550 504663 Homo sapiens Human secreted protein, SEQ ID NO: 8144. 84 4111 50 550 504663 Homo sapiens Human secreted protein, SEQ ID NO: 7384. 111 60 51 52 503267 Homo sapiens Human secreted protein, SEQ ID NO: 7384. 111 60 51 52 503267 Homo sapiens Human secreted protein, SEQ ID NO: 7384. 75 29 65 503267 Homo sapiens Human secreted protein, SEQ ID NO: 7384. 75 29 53 503267 Homo sapiens Human secreted protein, SEQ ID NO: 7384. 182 48 53 48 53 48 53 48 53 48 53 48 53 50326 Homo sapiens Human secreted protein, SEQ ID NO: 7384. 182 48 48 53 50326 Homo sapiens Human secreted protein, SEQ ID NO: 7384. 182 48 53 50326 Homo sapiens Human secreted protein, SEQ ID NO: 636. 644 75 53 64 64 64 64 64 64 64 6	344		_		253	/3
Section Sect	525	AF119851	Homo sapiens		162	57
Section Sect	526		Homo sapiens	Human secreted protein encoded by gene No. 47.	154	57
Section Sect	527	G02707	Homo sapiens		70	45
Signature			Homo sapiens	C8	1112	86
G04067			Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	84	45
G03267 Homo sapiens Human secreted protein, SEQ ID NO: 7348. 75 29 333 AF068286 Homo sapiens HUMAN secreted protein, SEQ ID NO: 7284. 182 48 48 47668286 Homo sapiens HDCMD38P 4861 100 10	530	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	111	60
G03203				Human secreted protein, SEQ ID NO: 8148.	92	
AF068286 Homo sapiens HDCMD38P 861 100				Human secreted protein, SEQ ID NO: 7348.		29
1975 1976			Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	182	48
					861	100
537 AF219232 Gallus gallus qin-induced kinase 206 53 538 AF135022 Homo sapiens mediator 128 1100 539 G03267 Homo sapiens Human secreted protein, SEQ ID NO: 7348. 141 59 540 AF016430 Caenorhabdit contains similarity to a BR-C/TIK domain selegans 853 39 541 AC003093 Homo sapiens OXYSTEROL-BINDING PROTEIN; 45% 408 66 542 M29487 Homo sapiens OXYSTEROL-BINDING PROTEIN; 45% 408 66 543 AF102530 Mus integrin alpha subunit precursor 517 81 543 AF102530 Mus olfactory receptor F3 327 73 544 Y73431 Homo sapiens Human secreted protein clone yb186_1 protein 386 100 545 AE004833 Pseudomonas probable TonB-dependent receptor 279 42 546 G03793 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 547					228	60
128 100					_ :: _ :	75
Social Company Soci						
AF016430 Caenorhabditi s clegams Contains similarity to a BR-C/TTK domain S53 39				· · · · · · · · · · · · · · · · · · ·	128	1
541 AC003093 Homo sapiens OXYSTEROL-BINDING PROTEIN; 45% similarity to P22059 (PID:g129308) 408 66 542 M29487 Homo sapiens integrin alpha subunit precursor 517 81 543 AF102530 Mus olfactory receptor F3 327 73 544 Y73431 Homo sapiens Human secreted protein clone yb186_1 protein 386 100 545 AE004833 Pseudomonas acruginosa probable TonB-dependent receptor 279 42 546 G03793 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 547 Y69192 Homo sapiens Human secreted protein sequence encoded by gene 43 SEQ ID NO: 166. 1772 67 548 Y91493 Homo sapiens Human secreted protein, SEQ ID NO: 5652. 777 99 550 AF044588 Homo sapiens protein regulating cytokinesis 1; PRC1 1953 83 551 Y29332 Homo sapiens Human secreted protein clone pe584_2 protein sequence. 194 94 552 X98330 Homo						
541 AC003093 Homo sapiens similarity to P22059 (PID:g129308) 408 66 542 M29487 Homo sapiens integrin alpha subunit precursor 517 81 543 AF102530 Mus musculus olfactory receptor F3 327 73 544 Y73431 Homo sapiens Human secreted protein clone yb186_1 protein sequence SEQ ID NO:84. 386 100 545 AE004833 Pseudomonas acruginosa Pseudomonas probable TonB-dependent receptor 279 42 546 G03793 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 547 Y69192 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 548 Y91493 Homo sapiens Human secreted protein sequence encoded by gene 43 SEQ ID NO: 166. 176 100 549 G01571 Homo sapiens Human secreted protein sequence encoded by gene 43 SEQ ID NO: 5652. 777 99 550 AF044588 Homo sapiens Human secreted protein clone pc584_2 protein sequence. 176 100 552 X98330			s elegans		853	39
542 M29487 Homo sapiens integrin alpha subunit precursor 517 81 543 AF102530 Mus olfactory receptor F3 327 73 544 Y73431 Homo sapiens Human secreted protein clone yb186_1 protein 386 100 545 AE004833 Pseudomonas acruginosa Probable TonB-dependent receptor 279 42 546 G03793 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 547 Y69192 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 548 Y91493 Homo sapiens Human secreted protein sequence encoded by gene 43 SEQ ID NO: 166. 176 100 549 G01571 Homo sapiens Human secreted protein, SEQ ID NO: 5652. 777 99 550 AF044588 Homo sapiens Protein regulating cytokinesis 1; PRC1 1953 88 551 Y29332 Homo sapiens Human secreted protein clone pe584_2 protein 1224 94 552 X98330 Homo sapiens Human secr	541	AC003093			408	66
543 AF102530 Mus musculus musculus olfactory receptor F3 327 73 544 Y73431 Homo sapiens sequence SEQ ID NO:84. 100 545 AE004833 Pseudomonas acruginosa 279 42 546 G03793 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 547 Y69192 Homo sapiens Homo sapiens Human secreted protein sequence encoded by Bene 43 SEQ ID NO: 166. 1772 67 548 Y91493 Homo sapiens Human secreted protein sequence encoded by gene 43 SEQ ID NO: 166. 176 100 549 G01571 Homo sapiens Human secreted protein, SEQ ID NO: 5652. 777 99 550 AF044588 Homo sapiens Human secreted protein, SEQ ID NO: 5652. 777 99 551 Y29332 Homo sapiens Protein regulating cytokinesis I; PRC1 1953 88 551 Y29330 Homo sapiens Protein receptor 2 24621 99 552 X98330 Homo sapiens Human UC Band #331 protein. 684 95 553 Y42782 Homo sapiens Protein receptor sub receptor sub receptor sub recep	542	M29487	Homo sapiens	integrin alpha subunit precursor	517	81
Museulus						
Sequence SEQ ID NO:84.		}			1 32.	1 "
545 AE004833 Pseudomonas acruginosa probable TonB-dependent receptor 279 42 546 G03793 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 547 Y69192 Homo sapiens A human monocyte-macrophage apolipoprotein Breceptor protein. 1772 67 548 Y91493 Homo sapiens Human secreted protein sequence encoded by gene 43 SEQ ID NO: 166. 176 100 549 G01571 Homo sapiens Human secreted protein, SEQ ID NO: 5652. 777 99 550 AF044588 Homo sapiens Human secreted protein clone pe584_2 protein sequence. 1953 88 551 Y29332 Homo sapiens Human secreted protein clone pe584_2 protein sequence. 1953 88 552 X98330 Homo sapiens Human receptor 2 24621 99 553 Y42782 Homo sapiens Human receptor 2 24621 99 554 AB025258 Mus granuphilin-a 501 41 555 A9010346 Homo sapiens Human TR-interacting prote	544	Y73431	Homo sapiens		386	100
Second Second	545	AE004833	Pseudomonas		279	42
546 G03793 Homo sapiens Human secreted protein, SEQ ID NO: 7874. 264 53 547 Y69192 Homo sapiens A human monocyte-macrophage apolipoprotein 1772 67 548 Y91493 Homo sapiens Human secreted protein sequence encoded by gene 43 SEQ ID NO: 166. 176 100 549 G01571 Homo sapiens Human secreted protein, SEQ ID NO: 5652. 777 99 550 AF044588 Homo sapiens Human secreted protein, SEQ ID NO: 5652. 777 99 551 Y29332 Homo sapiens Human secreted protein clone pe584_2 protein sequence. 1953 88 551 Y29332 Homo sapiens Human secreted protein clone pe584_2 protein sequence. 1224 94 552 X98330 Homo sapiens Human Secreted protein clone pe584_2 protein sequence. 1224 94 553 Y42782 Homo sapiens Human UC Band #331 protein. 684 95 554 AB025258 Mus musculus FRO1722 1468 100 555 AJ010346 Homo sapien		ì				1
547 Y69192 Homo sapiens B receptor protein. A human monocyte-macrophage apolipoprotein B receptor protein. 1772 67 548 Y91493 Homo sapiens Human secreted protein sequence encoded by gene 43 SEQ ID NO:166. 100 100 549 G01571 Homo sapiens Human secreted protein, SEQ ID NO: 5652. 777 99 550 AF044588 Homo sapiens Protein regulating cytokinesis 1; PRC1 1953 88 551 Y29332 Homo sapiens Protein regulating cytokinesis 1; PRC1 1953 88 551 Y29332 Homo sapiens Protein regulating cytokinesis 1; PRC1 1953 88 551 Y29332 Homo sapiens Protein regulating cytokinesis 1; PRC1 1953 88 551 Y29332 Homo sapiens Protein regulating cytokinesis 1; PRC1 1953 88 552 X98330 Homo sapiens Protein regulating cytokinesis 1; PRC1 1953 88 553 Y42782 Homo sapiens Protein regulating cytokinesis 1; PRC1 1224 94 554 AB025258 Mus Ruman secreted protein. 561 141 141 141 141 <td>546</td> <td>G03793</td> <td></td> <td>Human secreted protein, SEQ ID NO: 7874.</td> <td>264</td> <td>53</td>	546	G03793		Human secreted protein, SEQ ID NO: 7874.	264	53
Human secreted protein sequence encoded by gene 43 SEQ ID NO: 166. 100	547	Y69192	Homo sapiens	A human monocyte-macrophage apolipoprotein		
gene 43 SEQ ID NO:166.	549 .	V01402	Maria soriesa		100	1.00
550 AF044588 Homo sapiens protein regulating cytokinesis 1; PRC1 1953 88 551 Y29332 Homo sapiens Human secreted protein clone pe584_2 protein 1224 94 552 X98330 Homo sapiens ryanodine receptor 2 24621 99 553 Y42782 Homo sapiens Human UC Band #331 protein. 684 95 554 AB025258 Mus musculus granuphilin-a 501 41 555 AJ010346 Homo sapiens RING-H2 1468 100 556 W92388 Homo sapiens Human TR-interacting protein S239a. 538 92 557 AF119851 Homo sapiens PRO1722 175 59 558 AF117756 Homo sapiens thyroid hormone receptor-associated protein complex component TRAP150 183 32 559 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953 319 68 560 D86214 Mus Ca2+ dependent activator protein for secretion musculus 1010 93	J40	191493	Holno Sapiens	gene 43 SEQ ID NO:166.	176	100
550 AF044588 Homo sapiens protein regulating cytokinesis 1; PRC1 1953 88 551 Y29332 Homo sapiens Human secreted protein clone pe584_2 protein 1224 94 552 X98330 Homo sapiens ryanodine receptor 2 24621 99 553 Y42782 Homo sapiens Human UC Band #331 protein. 684 95 554 AB025258 Mus musculus granuphilin-a 501 41 555 AJ010346 Homo sapiens RING-H2 1468 100 556 W92388 Homo sapiens Human TR-interacting protein S239a. 538 92 557 AF119851 Homo sapiens PRO1722 175 59 558 AF117756 Homo sapiens thyroid hormone receptor-associated protein complex component TRAP150 183 32 559 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953 319 68 560 D86214 Mus Ca2+ dependent activator protein for secretion musculus 1010 93		G01571	Homo sapiens	Human secreted protein, SEQ ID NO: 5652.		99
Sequence Sequence	550	AF044588	Homo sapiens		1953	88
552 X98330 Homo sapiens ryanodine receptor 2 24621 99 553 Y42782 Homo sapiens Hurnan UC Band #331 protein. 684 95 554 AB025258 Mus musculus granuphilin-a 501 41 555 AJ010346 Homo sapiens RING-H2 1468 100 556 W92388 Homo sapiens Hurnan TR-interacting protein S239a. 538 92 557 AF119851 Homo sapiens PRO1722 175 59 558 AF117756 Homo sapiens thyroid hormone receptor-associated protein 183 32 559 G02872 Homo sapiens Hurnan secreted protein, SEQ ID NO: 6953. 319 68 560 D86214 Mus musculus Ca2+ dependent activator protein for secretion musculus 1010 93 561 AF187325 Canis familiaris melanoma antigen 287 55 562 AJ001981 Homo sapiens OXA1L 2512 99 563 Z17238 Ra	551	Y29332	Homo sapiens	Human secreted protein clone pe584_2 protein		94
553 Y42782 Homo sapiens Human UC Band #331 protein. 684 95 554 AB025258 Mus musculus granuphilin-a 501 41 555 AJ010346 Homo sapiens RING-H2 1468 100 556 W92388 Homo sapiens Human TR-interacting protein S239a. 538 92 557 AF119851 Homo sapiens PRO1722 175 59 558 AF117756 Homo sapiens thyroid hormone receptor-associated protein 183 32 559 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 319 68 560 D86214 Mus Ca2+ dependent activator protein for secretion musculus 1010 93 561 AF187325 Canis familiaris melanoma antigen 287 55 562 AJ001981 Homo sapiens OXA1L 2512 99 563 Z17238 Rattus norvegicus glutarnate receptor subtype delta-1 338 66 564 W30638	552	V09220	YYong polices			-
AB025258 Mus granuphilin-a 501 41						
Muse						
555 AJ010346 Homo sapiens RING-H2 1468 100 556 W92388 Homo sapiens Human TR-interacting protein S239a. 538 92 557 AF119851 Homo sapiens PRO1722 175 59 558 AF117756 Homo sapiens thyroid hormone receptor-associated protein complex component TRAP150 183 32 559 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 319 68 560 D86214 Mus Ca2+ dependent activator protein for secretion 1010 93 561 AF187325 Canis melanoma antigen 287 55 562 AJ001981 Homo sapiens OXA1L 2512 99 563 Z17238 Rattus norvegicus glutamate receptor subtype delta-1 norvegicus 338 66 564 W30638 Homo sapiens R33590_1 467 97 565 AC005620 Homo sapiens R33590_1 467 97 566 Y99358 Homo sapiens	JJ T	1. 12023236		gaaapiiini-a	301	41
S56 W92388 Homo sapiens Human TR-interacting protein S239a. S38 92	555	AJ010346		RING-H2	1468	100
557 AF119851 Homo sapiens PRO1722 175 59 558 AF117756 Homo sapiens thyroid hormone receptor-associated protein complex component TRAP150 183 32 559 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 319 68 560 D86214 Mus musculus Ca2+ dependent activator protein for secretion musculus 1010 93 561 AF187325 Canis familiaris melanoma antigen 287 55 562 AJ001981 Homo sapiens OXA1L 2512 99 563 Z17238 Rattus norvegicus glutamate receptor subtype delta-1 338 66 564 W30638 Homo sapiens Partial human 7-transmembrane receptor HAPO167 protein. 371 100 565 AC005620 Homo sapiens R33590_1 467 97 566 Y99358 Homo sapiens Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63. 1138 78 567 AL031177 Homo sapiens d/3889M15_3 (novel protein) 100					I	
558 AF117756 Homo sapiens thyroid hormone receptor-associated protein complex component TRAP150 183 32 559 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 319 68 560 D86214 Mus musculus Ca2+ dependent activator protein for secretion musculus 1010 93 561 AF187325 Canis familiaris melanoma antigen 287 55 562 AJ001981 Homo sapiens OXA1L 2512 99 563 Z17238 Rattus norvegicus glutamate receptor subtype delta-1 morvegicus 338 66 564 W30638 Homo sapiens Partial human 7-transmembrane receptor HAPO167 protein. 371 100 565 AC005620 Homo sapiens R33590_1 467 97 566 Y99358 Homo sapiens Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63. 1138 78 567 AL031177 Homo sapiens dJ889M15.3 (novel protein) 1002 58						
Complex component TRAP150 Section Sectio						
559 G02872 Homo sapiens Human secreted protein, SEQ ID NO: 6953. 319 68 560 D86214 Mus musculus Ca2+ dependent activator protein for secretion musculus 1010 93 561 AF187325 Canis familiaris melanoma antigen 287 55 562 AJ001981 Homo sapiens OXA1L 2512 99 563 Z17238 Rattus norvegicus glutamate receptor subtype delta-1 glutamate receptor subtype delta-1 human 7-transmembrane receptor HAPO167 protein. 371 100 564 W30638 Homo sapiens R33590_1 467 97 565 AC005620 Homo sapiens Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63. 1138 78 567 AL031177 Homo sapiens dJ889M15.3 (novel protein) 1002 58		1	IIIII Suprema		103	32
560 D86214 Mus musculus Ca2+ dependent activator protein for secretion musculus 1010 93 561 AF187325 Canis familiaris melanoma antigen 287 55 562 AJ001981 Homo sapiens OXA1L 2512 99 563 Z17238 Rattus norvegicus glutamate receptor subtype delta-1 338 66 564 W30638 Homo sapiens Partial human 7-transmembrane receptor HAPO167 protein. 371 100 565 AC005620 Homo sapiens R33590_1 467 97 566 Y99358 Homo sapiens Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63. 1138 78 567 AL031177 Homo sapiens d3889M15.3 (novel protein) 1002 58	559	G02872	Homo saniens		319	68
musculus						
561 AF187325 Canis familiaris melanoma antigen 287 55 562 AJ001981 Homo sapiens OXA1L 2512 99 563 Z17238 Rattus norvegicus glutamate receptor subtype delta-1 338 66 564 W30638 Homo sapiens Partial human 7-transmembrane receptor HAPO167 protein. 371 100 565 AC005620 Homo sapiens R33590_1 467 97 566 Y99358 Homo sapiens Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63. 1138 78 567 AL031177 Homo sapiens d3889M15.3 (novel protein) 1002 58			1 '	Proceedings brocom for postotion	1	"
familiaris 562 AJ001981 Homo sapiens OXA1L 2512 99 563 Z17238 Rattus norvegicus glutamate receptor subtype delta-1 338 66 564 W30638 Homo sapiens Partial human 7-transmembrane receptor HAPO167 protein. 371 100 565 AC005620 Homo sapiens R33590_1 467 97 566 Y99358 Homo sapiens Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63. 1138 78 567 AL031177 Homo sapiens d3889M15.3 (novel protein) 1002 58	561	AF187325		melanoma antigen	287	55
562 AJ001981 Homo sapiens OXA1L 2512 99 563 Z17238 Rattus norvegicus glutamate receptor subtype delta-1 338 66 564 W30638 Homo sapiens Partial human 7-transmembrane receptor HAPO167 protein. 371 100 565 AC005620 Homo sapiens R33590_1 467 97 566 Y99358 Homo sapiens Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63. 1138 78 567 AL031177 Homo sapiens d3889M15.3 (novel protein) 1002 58		1	1			
563 Z17238 Rattus norvegicus glutamate receptor subtype delta-1 338 66 564 W30638 Homo sapiens Partial human 7-transmembrane receptor HAPO167 protein. 371 100 565 AC005620 Homo sapiens R33590_1 467 97 566 Y99358 Homo sapiens Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63. 1138 78 567 AL031177 Homo sapiens d3889M15.3 (novel protein) 1002 58	562	AJ001981	Homo sapiens	OXAIL	2512	99
norvegicus 100	563					
HAPO167 protein. Section 2016 HAPO167 protein. HAPO167 protein			norvegicus	• • •		
565 AC005620 Homo sapiens R33590_1 467 97 566 Y99358 Homo sapiens Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63. 1138 78 567 AL031177 Homo sapiens dJ889M15.3 (novel protein) 1002 58	564	W30638	Homo sapiens		371	100
566 Y99358 Homo sapiens Human PRO1772 (UNQ834) amino acid 1138 78 sequence SEQ ID NO:63. sequence SEQ ID NO:63. 1002 58	ECE	10000000	17			<u> </u>
sequence SEQ ID NO:63.						
100Z 30		<u> </u>		sequence SEQ ID NO:63.		
368 AF151043 Homo sapiens HSPC209 798 100						
	208	AF151043	Homo sapiens	HSPC209	798	100

SEO	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	1.0.	}		Score	
569	AF097518	Homo sapiens	liver-specific transporter	231	100
570	AB035698	Homo sapiens	Misshapen/NIK-related kinase MINK-1	1532	100
571	Y07096	Homo sapiens	Colon cancer associated antigen precursor	1064	100
3/1	107096	Homo sapiens	J .	1004	100
		l	sequence.	73.5	1 55
572	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	735	55
573	Y66639	Homo sapiens	Membrane-bound protein PRO290.	254	45
574	AB037108	Homo sapiens	seven transmembrane domain orphan receptor	1883	99
575	D43949	Homo sapiens	This gene is novel.	836	100
576	Y48596	Homo sapiens	Human breast tumour-associated protein 57.	108	50
577	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	141	75
578	R95913	Homo sapiens	Neural thread protein.	140	65
579	AK025116	Homo sapiens	unnamed protein product	201	70
580	Y86473	Homo sapiens	Human gene 52-encoded protein fragment, SEQ	77	70
		- zem supreme	ID NO:388.		
581	AF196779	Homo sapiens	JM10 protein	450	100
582	AF188706	Homo sapiens	g20 protein	330	98
583	AB030234	Canis	D4 dopamine receptor	64	56
دەد	7.5030234	familiaris	D4 dobatime recebior	U ⁴⁴	20
604	00000		II	245	90
584	G02621	Homo sapiens	Human secreted protein, SEQ ID NO: 6702.	345	1 2-7
585	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	268	85
			Antigen)		<u> </u>
586	Y30819	Homo sapiens	Human secreted protein encoded from gene 9.	235	35
587	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	132	56
588	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	182	79
589	AF235017	Mus	2P1 protein	764	80
	i	musculus			
590	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	329	81
			HPMBQ32.		
591	Y30709	Homo sapiens	Amino acid sequence of a human secreted	110	43
	120,00	1101110 54714110	protein.		1
592	Y53875	Homo sapiens	A human seven transmembrane signal transducer	1369	92
392	133073	Troing sabiens		1309	1 32
593	Y53051	170-1-0-1-0-	polypeptide.	1112	97
393	1 23031	Homo sapiens	Human secreted protein clone dd119_4 protein	1112	91
504	1,05650	 _	sequence SEQ ID NO:108.	762	70
594	Y27658	Homo sapiens	Human secreted protein encoded by gene No. 92.	763	79
595	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	156	58
596	AF151110	Mus	COP1 protein	2215	95
		musculus			<u> </u>
597	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	157	65
598 ⁻	AF192499	Mus	putative secreted protein ZSIG37	143	40
		musculus			
599	AF119855	Homo sapiens	PRO1847	236	76
600	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	212	73
601	Y00295	Homo sapiens	Human secreted protein encoded by gene 38.	567	88
602	AF184971	Homo sapiens	class II cytokine receptor ZCYTOR7	2015	74
603	AF061936	Homo sapiens	diacylglycerol kinase iota	773	96
604	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	1333	93
504	ALU30020	1101110 sapiens		1333	1 ,3
605	AD022106	Tions as '	Antigen)	2015	100
605	AB033106	Homo sapiens	KIAA1280 protein	3915	100
606	X75756	Homo sapiens	protein kinase C mu	3916	99
607	D86983	Homo sapiens	similar to D.melanogaster peroxidasin(U11052)	5758	99
608	W69341	Homo sapiens	Secreted protein of clone CG279_1.	1377	99
609	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	339	82
		1	HPMBQ32.		1
610	Y27868	Homo sapiens	Human secreted protein encoded by gene No.	116	62
			107.		
611	AF202636	Homo sapiens	angiopoietin-like protein PP1158	2164	100
612	AF090944	Homo sapiens	PRO0663	218	82
613	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	195	59
013	1 02033	Trouto sapiens		153	ا ا
<u> </u>	1407052	Pottura	clone HTDAD22.	450	104
614	M87053	Rattus	lens membrane protein	450	84
	1	norvegicus			
615	AC004232 G01984	Homo sapiens	FPM315	163	37
616		Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	205	79

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
617	Y91524	Homo sapiens	Human secreted protein sequence encoded by gene 74 SEQ ID NO:197.	821	99
618	AJ245621	Homo sapiens	CTL2 protein	2258	99
619	Y76198	Homo sapiens	Human secreted protein encoded by gene 75.	108	64
620	AF067864	Homo sapiens	transferrin receptor 2 alpha	3922	94
621	D90721	Escherichia coli	Transmembrane protein dppC	573	90
622	W75858	Homo sapiens	Human secretory protein of clone CS752-3.	730	100
623	Y94982	Homo sapiens	Human secreted protein vb12_1, SEQ ID NO:4.	733	100
624	AF034745	Mus musculus	LNXp80	637	83
625	U42580	Paramecium bursaria Chlorella virus I	Pro-rich, IPPPNMSLPLS (3x)	94	46
626	U79260	Homo sapiens	unknown	194	70
627	R95913	Homo sapiens	Neural thread protein.	99	50
628	G03450	Homo sapiens	Human secreted protein, SEQ ID NO: 7531.	427	100
629	Y36281	Homo sapiens	Human secreted protein encoded by gene 58.	590	100
630	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	165	76
631	G02139	Homo sapiens	Human secreted protein, SEQ ID NO: 6220.	268	96
632	U16996	Homo sapiens	protein tyrosine posphatase	351	80
633	AF121857	Homo sapiens	sorting nexin 7	2019	100
634	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899	340	77
635	Y07090	Homo sapiens	Renal cancer associated antigen precursor sequence.	277	64
636	AB013382	Homo sapiens	DUSP6	414	76
637	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	315	71
638	M95762	Rattus norvegicus	GABA transporter	924	89
639	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	219	60
640	Y01400	Homo sapiens	Secreted protein encoded by gene 18 clone HNHFO29.	137	79
641	AC008075	Arabidopsis thaliana	F24J5.4	121	33
642	W74824	Homo sapiens	Human secreted protein encoded by gene 96 clone HAQBK61.	615	62
643	AB015982	Homo sapiens	serine/threonine kinase	485	98
644	Y25806	Homo sapiens	Human secreted protein fragment encoded from gene 23.	162	46
645	AF122904	Homo sapiens	membrane protein DAP10	474	100
646	AF233323	Homo sapiens	Fas-associated phosphatase-1	200	38
647	W48804	Homo sapiens	Homo sapiens clone BK158_1 protein.	1203	99
648	AF257330	Homo sapiens	COBW-like protein	1440	98
649	Y36203	Homo sapiens	Human secreted protein #75.	233	73
650	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	173	78
651	Y32199	Homo sapiens	Human receptor molecule (REC) encoded by Incyte clone 2022379.	1012	100
652	AB032909	Hylobates agilis	dopamine receptor D4	122	32
653	AK021848	Homo sapiens	unnamed protein product	186	69
654	W73411	Homo sapiens	Human secreted protein encoded by Gene No. 15.	57	37
655	L22455	Rattus norvegicus	mu opioid receptor	116	34
656	G03112	Homo sapiens	Human secreted protein, SEQ ID NO: 7193.	110	45
657	G02345	Homo sapiens	Human secreted protein, SEQ ID NO: 6426.	459	97
658	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	291	75
659	G02832	Homo sapiens	Human secreted protein, SEQ ID NO: 6913.	134	65
660	Y91423	Homo sapiens	Human secreted protein sequence encoded by gene 11 SEQ ID NO:144.	333	96

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
661	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	168	68 .
662	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	375	43
002	133660	Homo sapiens	designated HSCOP-6.		1
663	W75771	Homo sapiens	Human GTP binding protein APD08.	629	100
664	AL096770	Homo sapiens	bA150A6.2 (novel 7 transmembrane receptor	480	55
			(rhodopsin family) (olfactory receptor like) protein (hs6M1-21))		
665	AB037734	Homo sapiens	KIAA1313 protein	978	96
666	W82841	Homo sapiens	Human cerebral protein-1.	192	84
667	W82841	Homo sapiens	Human cerebral protein-1.	182	87
668	AB030184	Mus musculus	contains transmembrane (TM) region and ATP binding region	757	68
669	AB032919	Hylobates muelleri	dopamine receptor D4	85	37
670	AF107295	Rattus norvegicus	outer membrane protein	746	81
671	Z33642	Homo sapiens	leukocyte surface protein	394	93
672	W85608	Homo sapiens	Secreted protein clone du410_5.	261	91
673	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	106	48
674	AL035587	Homo sapiens	dJ475N16.4 (KIAA0240)	2388	99
675	Y59668	Homo sapiens	Secreted protein 108-005-5-0-C1-FL.	1134	53
676	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	174	74
677	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor; PDPr	1013	95
678	L11625	Mus musculus	receptor protein-tyrosine kinase	545	96
679	AL031427	Homo sapiens	dJ167A19.3 (novel protein)	745	100
680	AJ133430	Mus musculus	olfactory receptor	528	77
681	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	179	70
682	G02332	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	336	76
683	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	118	100
684	U43360	Peromyscus maniculatus	reverse transcriptase	100	37
685	G00885	Homo sapiens	Human secreted protein, SEQ ID NO: 4966.	162	60
686	AK001518	Homo sapiens	unnamed protein product	590	100
687	G01982	Homo sapiens	Human secreted protein, SEQ ID NO: 6063.	718	100
688	Y92241	Homo sapiens	Human cancer associated antigen precursor (MO-REN-46).	2405	99
689	AC024792	Caenorhabditi s elegans	contains similarity to TR:P78316	423	36
690	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	183	81
691	Y56514	Homo sapiens	Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence.	180	88
692	Y27795	Homo sapiens	Human secreted protein encoded by gene No. 79.	1539	99
693	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	428	98
694	U12465	Homo sapiens	ribosomal protein L35	308	89
695	Y45272	Homo sapiens	Human secreted protein encoded from gene 16.	1517	99
696	AF191838	Homo sapiens	TANK binding kinase TBK1	1242	98
697	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	275	75
698	Y87280	Homo sapiens	Human signal peptide containing protein HSPP- .57 SEQ ID NO:57.	576	90
699	Y97999	Homo sapiens	Human SCAD family molecule HSFM-1, SEQ ID NO:1.	729	99
700	AJ006701	Homo sapiens	putative serine/threonine protein kinase	610	79
701	AF209198	Homo sapiens	zinc finger protein 277	2357	100
702	AJ298841	Mus musculus	torsinA protein	709	45
703	AK021729	Homo sapiens	unnamed protein product	622	98
704	Z46787	Caenorhabditi s elegans	similar to Glutaredoxin, Zinc finger, C3HC4 type (RING finger)	920	51
705	G02882	Homo sapiens	Human secreted protein, SEQ ID NO: 6963.	589	98

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
706	G02501	Homo sapiens	Human secreted protein, SEQ ID NO: 6582.	125	58
707	R95326	Homo sapiens	Tumor necrosis factor receptor 1 death domain ligand (clone 2DD).	121	95
708	G03002	Homo sapiens	Human secreted protein, SEQ ID NO: 7083.	125	39
709	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	516	98
710	M63577	Saccharomyc es cerevisiae	SFP1	131	59
711	AB026291	Rattus norvegicus	acetoacetyl-CoA synthetase	467	85
712	D21211	Homo sapiens	protein tyrosine phosphatase (PTP-BAS, type 3)	368	44
713	AF044033	Marmota marmota	olfactory receptor	615	83
714	G03561	Homo sapiens	Human secreted protein, SEQ ID NO: 7642.	251	100
715	AB033062	Homo sapiens	KIAA1236 protein	1380	100
716	G00577	Homo sapiens	Human secreted protein, SEQ ID NO: 4658.	80	73
717	Y96864	Homo sapiens	SEQ. ID. 37 from WO0034474.	835	99
718 719	AJ243396 U47334	Homo sapiens Homo sapiens	voltage-gated sodium channel beta-3 subunit similar to chicken gamma aminobutyric acid	234 578	100 99
720	AB020598	Homo sapiens	receptor beta4 subunit peptide transporter 3	1096	100
721	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	570	74
722	J05046	Homo sapiens	insulin receptor-related receptor	6787	100
723	AF001958	Ambystoma tigrinum	electrogenic Na+ bicarbonate cotransporter; NBC	111	41
724	AF127084	Mus musculus	semaphorin cytoplasmic domain-associated protein 3A	5253	94
725	X54673	Homo sapiens	GABA transporter	3114	99
726	AF016191	Rattus norvegicus	potassium channel	370	100
727	AB029559	Rattus . norvegicus	BATI	139	35
728	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	2186	97
729 730	AJ011415 Z93096	Homo sapiens Homo sapiens	plexin-B1/SEP receptor bK390B3.1 (manic fringe (Drosophila)	729 142	56 68
731	Z10062	Homo sapiens	homolog) cDNA encoding a human vanilloid receptor homologue Vanilrep1.	675	99
732	AF161382	Homo sapiens	HSPC264	492	94
733	AB029033	Homo sapiens	KIAA1110 protein	3826	99
734	AE000493	Escherichia coli	putative transport protein	592	97
735	AL033379	Homo sapiens	dJ417O22.2 (novel 7 transmembrane receptor (rhodopsin family) protein similar to high- affinity lysophosphatidic acid receptor homolog)	2173	99
736	AF132599	Homo sapiens	RANTES factor of late activated T lymphocytes- 1	245	56
737	X55019	Homo sapiens	acetylcholine receptor delta subunit	883	99.
738	X91906	Homo sapiens	voltage-gated chloride ion channel	1978	100
739 740	AB026116 D00570	Homo sapiens Mus musculus	organic anion transporter 4 open reading frame (196 AA)	1444 83	98 24
741	W03626	Homo sapiens	Human thyrotropin GPR N-terminal sequence.	118	40
742	U66059	Homo sapiens	V_segment translation product	614	100
743	AF119815	Homo sapiens	G-protein-coupled receptor	2751	99
744	X16663	Homo sapiens	haematopoietic lineage cell protein (AA 1-486)	148	93
745	W67838	Homo sapiens	Human secreted protein encoded by gene 32 clone HLTCJ63.	448	95
746	W57260	Homo sapiens	Human semaphorin Y.	2414 .	100
747	W21578	Homo sapiens	Alzheimer's disease protein encoded by DNA from plasmid pGCS2232.	968	65
748	Y94935	Homo sapiens	Human secreted protein clone yd218_1 protein sequence SEQ ID NO:76.	622	100
749	AL022238	Homo sapiens	dJ1042K10.5 (novel protein)	314	85
750	G03889	Homo sapiens	Human secreted protein, SEQ ID NO: 7970.	391	87

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
751	AB025258	Mus musculus	granuphilin-a	773	41
7.50	V63206		17	900	99
752	Y52386	Homo sapiens	Human transmembrane protein HP02000.	2527	99
753	Y48586	Homo sapiens	Human breast tumour-associated protein 47.		
754	AJ272207	Homo sapiens	putative G protein-coupled receptor 92	694	100
755	M85183	Rattus norvegicus	vasopressin receptor	979	68
756	AF190501	Homo sapiens	leucine-rich repeat-containing G protein-coupled receptor 6	388	71
757	Y02692	Homo sapiens	Human secreted protein encoded by gene 43 clone HTADX17.	461	87
758	Z22535	Homo sapiens	ALK-3	439	98
759	R04932	Homo sapiens	Interferon-gamma receptor segment from clone 39 responsible for binding the target.	564	97
760	W74902	Homo sapiens	Human secreted protein encoded by gene 175 clone HE8BI92.	1217	99
761	G03706	Homo sapiens	Human secreted protein, SEQ ID NO: 7787.	223	88
762	AB020676	Homo sapiens	KIAA0869 protein	4433	99
763	AK026992	Homo sapiens	unnamed protein product	2285	99
764	AF173358	Homo sapiens	glucocorticoid receptor AF-1 coactivator-1	573	100
765	AF268066	Mus musculus	netrin 4	2019	89
766	Y48585	Homo sapiens	Human breast tumour-associated protein 46.	1169	89
767	AF230378	Mus musculus	interleukin-1 delta	309	45
768	AF121975	Mus musculus	odorant receptor S18	268	62
769	AB008515	Homo sapiens	RanBPM	611	57
770	Y09945	Rattus norvegicus	putative integral membrane transport protein	458	50
771	AF226731	Homo sapiens	AD026	688	99
772	Y27132	Homo sapiens	Human glioblastoma-derived polypeptide (clone OA004FG).	1384	100
773	X87832	Homo sapiens	NOV/plexin-A1 protein	1821	98
774	AB025258	Mus musculus	granuphilin-a	500	41
775	AF125101	Homo sapiens	HSPC040 protein	232	93
776	G02815	Homo sapiens	Human secreted protein, SEQ ID NO: 6896.	314	95
777	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	191	68
778	R03301	Homo sapiens	Sequence of pre-human atrial natriuretic peptide.	213	45
779	AL357374	Homo sapiens	bA353C18.2 (novel protein)	232	100
780	AF100346	Homo sapiens	neuronal voltage gated calcium channel gamma- 3 subunit	1434 .	89
781	Y19566	Homo sapiens	Amino acid sequence of a human secreted protein.	103	52
782	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	1098	93
783	AF084464	Rattus norvegicus	GTP-binding protein REM2	141	30
784	W49042	Homo sapiens	Human low density lipoprotein binding protein LBP-3.	2693	99
785	AF238381	Homo sapiens	PTOV1	1904	91
786	Y91870	Homo sapiens	Human apoptosis related protein.	547	100
787	Y71062	Homo sapiens	Human membrane transport protein, MTRP-7.	1062	94
788	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	8684	98
789	AL049569	Homo sapiens	dJ37C10.3 (novel ATPase)	2848	96
790	AF151848	Homo sapiens	CGI-90 protein	745	96
791	Y08639	Homo sapiens	nuclear orphan receptor ROR-beta	1421	95
791 <u> </u>	Y41706	Homo sapiens	Human PRO381 protein sequence.	644	99
793	AF121228	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP95	1037	100
794	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	124	62
794 795	Y69384	Homo sapiens	Amino acid sequence of a 14274 receptor protein.	119	100

SEQ	Accession	Species	Description	Smith-	%
ID `	No.			Waterman	Identity
NO:	<u> </u>		·	Score	
797	AF258340	Homo sapiens	hepatocellular carcinoma-associated antigen 112	1151	99
798	AF159615	Homo sapiens	FGF receptor activating protein 1	461	98
799	Y59863	Homo sapiens	Human normal uterus tissue derived protein 26.	797	99
800	W70459	Homo sapiens	Human T1-receptor ligand III splice variant 2.	572	92
801	L00073	Homo sapiens	renin	1913	93
802	P92219	Homo sapiens	CRI protein.	11963	97
	<u> </u>	(human)			
803	X15357	Homo sapiens	ANP-A receptor preprotein (AA -32 to 1029)	5199	98
804	W64473	Homo sapiens	Human secreted protein from clone EC172_1.	4018	95
805	AJ243874	Homo sapiens	oligophrenin-4	2067	100
806	G01731	Homo sapiens	Human secreted protein, SEQ ID NO: 5812.	284	100
807	Z24680	Homo sapiens	garp	1562	83
808	AF171669	Homo sapiens	glycoprotein-associated amino acid transporter LAT2	1364	90
809	W70321	Homo sapiens	Secreted protein CC198_1.	1154	96
810	W74843	Homo sapiens	Human secreted protein encoded by gene 115 clone HOVBA03.	855	99
811	AF108831	Homo sapiens	K:Cl cotransporter 3	4561	100
812	AF092135	Homo sapiens	PTD014	862	100
813	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10	784	100
			encoded by GenBank Accession Number	100	100
814	G01563	Homo sapiens	Human secreted protein, SEQ ID NO: 5644.	330	100
815	AF051151	Homo sapiens	Toll/interleukin-1 receptor-like protein 3	3850	99
816	W95630	Homo sapiens	Homo sapiens secreted protein gene clone	358	100
817	G01082		gn114_1.		
818	AF151800	Homo sapiens	Human secreted protein, SEQ ID NO: 5163.	549	100
		Homo sapiens	CGI-41 protein	1106	95
819 820	L00352	Homo sapiens	low density lipoprotein receptor	3980	100
	X04434	Homo sapiens	IGF-I receptor	5832	99
821	G03844	Homo sapiens	Human secreted protein, SEQ ID NO: 7925.	572	100
822	AF212220	Homo sapiens	TERA	396	48
823	Y50125	Homo sapiens	Human glycophosphatidylinositol-anchored protein GPI-122.	4897	99
824	AF156778	Homo sapiens	ASB-3 protein	2675	98
825	AF096322	Homo sapiens	neuronal voltage-gated calcium channel gamma- 2 subunit	1105	100
826	Y07972	Homo sapiens	Human secreted protein fragment #2 encoded from gene 28.	1540	100
827	AB032013	Homo sapiens	potassium channel Kv8.1	2435	95
828	Y13620	Homo sapiens	BCL9	5284	96
829	Y91474	Homo sapiens	Human secreted protein sequence encoded by gene 24 SEQ ID NO:147.	541	98
830	X54232	Homo sapiens	glypican	1625	87
831	X14830	Homo sapiens	acetylcholine receptor beta-subunit preprotein	2540	100
832	Y71262	Homo sapiens	Human chondromodulin-like protein, Zchm1.	1002	100
833	G03873	Homo sapiens	Human secreted protein, SEQ ID NO: 7954.	638	96
834	AC003030	Homo sapiens	R29828_1	1389	93
835	Y38422	Homo sapiens	Human secreted protein.	964	87
836	U41557	Caenorhabditi s elegans	glycine-rich	85	36
837	AL121889	Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803))	998	75
838	AJ011415	Homo sapiens	plexin-B1/SEP receptor	1580	60
839	W80398	Homo sapiens	A secreted protein encoded by clone cw1543 3.	1105	67
840	G00862	Homo sapiens	Human secreted protein, SEQ ID NO: 4943.	255	92
841	G02650	Homo sapiens	Human secreted protein, SEQ ID NO: 6731.	644	97
842	AF036717	Homo sapiens	FGFR signalling adaptor SNT-1	2629	99
843	Y73446	Homo sapiens	Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114.	1089	100
844	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	357	60
845	AF151810	Homo sapiens	CGI-52 protein	1443	69 88
846	X83378	Homo sapiens	putative chloride channel	1620	
	AC004883	L ATOMO Sapicità	similar to general transcription factor 2I; similar	655	99

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
		<u></u>	to AF038969 (PID:g2827207)		
848	X99886	Homo sapiens	monocyte chemotactic protein-2	160	76
849	AC005587	Homo sapiens	similar to mouse olfactory receptor 13; similar to P34984 (PID:g464305)	963	98
850	AB038237	Homo sapiens	G protein-coupled receptor C5L2	1767	100
851	AF124490	Homo sapiens	ARF GTPase-activating protein GIT1	3415	98
852	Y86217	Homo sapiens	Human secreted protein HWHGU54, SEQ ID NO:132.	1189	99
853	AF224741	Homo sapiens	chloride channel protein 7	3748	99
854	X17094	Homo sapiens	furin (AA 1-794)	3550	99
855	W78245	Homo sapiens	Fragment of human secreted protein encoded by gene 19.	1245	99
856	R97569	Homo sapiens	Interleukin-2 receptor associated protein p43.	1926	100
857	Y41765	Homo sapiens	Human PRO1083 protein sequence.	3211	99
858	AF057306	Homo sapiens	transmembrane proteolipid	481	84
859	AK025116	Homo sapiens	unnamed protein product	374	69
860	Y41312	Homo sapiens	Human secreted protein encoded by gene 5 clone HLDRM43.	824	100
862	Y25776	Homo sapiens	Human secreted protein encoded from gene 66.	895	99
863	Y74188	Homo sapiens	Human prostate tumor EST fragment derived protein #375.	96	30
864	AF167473	Homo sapiens	heme-binding protein	870	99
865	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	211	67
866	X54870	Homo sapiens	Type II integral membrane protein	1201	100
867	G00700	Homo sapiens	Human secreted protein, SEQ ID NO: 4781.	640	99
868	Y07894	Homo sapiens	Human secreted protein fragment encoded from gene 43.	388	88
869	J00123	Homo sapiens	preproenkephalin (1349	95
870	Y91632	Homo sapiens	Human secreted protein sequence encoded by gene 25 SEQ ID NO:305.	1048	98
871	L04311	Homo sapiens	GABA-alpha receptor beta-3 subunit	237	93
872 ·	Y29988	Homo sapiens	Human cytokine family member EF-7 protein.	960	94
873	AF161382	Homo sapiens	HSPC264	1124	99
874	G03412	Homo sapiens	Human secreted protein, SEQ ID NO: 7493.	464	100
875	Y27572	Homo sapiens	Human secreted protein encoded by gene No. 6.	573	96
876	M15530	Homo sapiens	B-cell growth factor	171	56
877	W63681	Homo sapiens	Human secreted protein 1.	1652	99
878	L27867	Rattus norvegicus	neurexophilin	1448	98
879	Y10835	Homo sapiens	Amino acid sequence of a human secreted protein.	321	100
880	W88991	Homo sapiens	Polypeptide fragment encoded by gene 144.	936	100
881	AF118670	Homo sapiens	orphan G protein-coupled receptor	1971	100
882	AF208865	Homo sapiens	EDRF	528	100
883	Y18462	Homo sapiens	cathepsin L	209	72
884	Y94950	Homo sapiens	Human secreted protein clone dh1073_12 protein sequence SEQ ID NO:106.	348	100
885	AF070661	Homo sapiens	HSPC005	404	100
886	Y04315	Homo sapiens	Human secreted protein encoded by gene 23.	385	100
887 888	X92744 Y22496	Homo sapiens Homo sapiens	hBD-1 Human secreted protein sequence clone	375 994	94
000	V41202	Home	cn621_8.	4505	
889 890	Y41293 G03714	Homo sapiens Homo sapiens	Human soluble protein ZTMPO-1. Human secreted protein, SEQ ID NO: 7795.	4595 147	63
891	AF208856	Homo sapiens	BM-014	1012	99
892	U29195	Homo sapiens	neuronal pentraxin II	2002	98
893	X68149	Homo sapiens	Burkitt lymphoma receptor 1	1953	100
894	Y94914	Homo sapiens	Human secreted protein clone pw337_6 protein sequence SEQ ID NO:34.	537	100
895	W61630	Homo sapiens	Clone HNFGW06 of EGFR receptor family.	326	63
896	M24110	Homo sapiens	G0S19-2 peptide precursor	481	100
897	Z68747	Homo sapiens	imogen 38	2018	99
898	AF186112	Homo sapiens	neurokinin B-like protein ZNEUROK1	619	100
				U. U. U. U. U. U. U. U. U. U. U. U. U. U	

SEQ	Accession	Species	Description	Smith-	%
ID	No.	1	•	Waterman	Identity
NO:	<u> </u>			Score	
900	P60657	Homo sapiens	Sequence of human lipocortin.	1835	100
901	M27288	Homo sapiens	oncostatin M	1297	99
902	W85737	Homo sapiens	Polypeptide with transmembrane domain.	749	100
903	G01349	Homo sapiens	Human secreted protein, SEQ ID NO: 5430.	650	99
904	Y00261	Homo sapiens	Human secreted protein encoded by gene 4.	1133	99
905	AF039688	Homo sapiens	antigen NY-CO-3	771	99
906	AB007836	Homo sapiens	Hic-5	2544	100
907	AB017507	Homo sapiens	Apg12	224	100
908	AK000056	Homo sapiens	unnamed protein product	1537	98
909	¥86299	Homo sapiens	Human secreted protein HFOXB55, SEQ ID NO:214.	427	100
910	AF231023	Homo sapiens	protocadherin Flamingo I	7393	99
911	Y14134	Homo sapiens	Vascular endothelial cell growth inhibitor beta protein sequence.	1319	100
912	Z90420	Homo sapiens	Human GDF-3 (hGDF-3) polypeptide encoding cDNA.	1950	100
913	Y19757	Homo sapiens	SEQ ID NO 475 from WO9922243.	1361	100
914	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	112	48
915	U14971	Homo sapiens	ribosomal protein S9	886	90
916	AF172854	Homo sapiens	cardiotrophin-like cytokine CLC	1204	99
917	AC005525	Homo sapiens	F22162_1	1963	100
918	AF166350	Homo sapiens	ST7 protein	4711	99
919	Y87285	Homo sapiens	Human signal peptide containing protein HSPP- 62 SEQ ID NO:62.	430	100
920	Y36131	Homo sapiens	Human secreted protein #3.	465	88
921	AF193766	Homo sapiens	cytokine-like protein C17	724	100
922	Y95013	Homo sapiens	Human secreted protein vc48_1, SEQ ID NO:66.	357	100
923	X75208	Homo sapiens	protein tyrosine kinase-receptor	5256	100
924	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	813	98
925	AB039886	Homo sapiens	down-regulated in gastric cancer	785	78
926	G03368	Homo sapiens	Human secreted protein, SEQ ID NO: 7449.	55	50
927	Y48606	Homo sapiens	Human breast tumour-associated protein 67.	539	100
928	Y36151	Homo sapiens	Human secreted protein #23.	668	100
929	AF110399	Homo sapiens	elongation factor Ts	1666	100
930	AF210317	Homo sapiens	facilitative glucose transporter family member GLUT9	2763	99
931	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931	100
932	G01959	Homo sapiens	Human secreted protein, SEQ ID NO: 6040.	274	100
933	U47924	Homo sapiens	B-cell receptor associated protein	1469	100
934	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	529	93
935	AB039371	Homo sapiens	mitochondrial ABC transporter 3	196	63
936	X56385	Canis	rab8	1064	100
		familiaris			1
937	B08906	Homo sapiens	Human secreted protein sequence encoded by gene 16 SEQ ID NO:63.	117	44
938	M13692	Homo sapiens	alpha-1 acid glycoprotein precursor	1064	99
939	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	515	42
940	Y16630	Homo sapiens	Human Putative Adrenomedullin Receptor (PAR).	1904	99
941	AC005102	Homo sapiens	small inducible cytokine subfamily A member 24	627	99
942	M12886	Homo sapiens	T-cell receptor beta chain	1289	81
943	AF226046	Homo sapiens	GK003	1049	98
944	Y36078	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 463.	667	100
945	M22877	Homo sapiens	cytochrome c	565	100
946	W67869	Homo sapiens	Human secreted protein encoded by gene 63 clone HHGDB72.	551	93
947	W67859	Homo sapiens	Human secreted protein encoded by gene 53 clone HBMCL41.	283	100
948	W85726	Homo sapiens	Novel protein (Clone BG33_7).	789	100
949	AJ242015	Homo sapiens	eMDC II protein	4236	100
		Troum on publicity	at brotom	7230	1 100

ID `		Species	Description	Smith-	%
	No.			Waterman	Identity
NO:	, i			Score	
951	AF110645	Homo sapiens	candidate tumor suppressor p33 ING1 homolog	1314	100
952	Y36111	Homo sapiens	Extended human secreted protein sequence, SEQ	402	70
		•	ID NO. 496.		
953	AB012109	Homo sapiens	APC10	990	100
954	AF246221	Homo sapiens	transmembrane protein BRI	1405	100
955	AF054986	Homo sapiens	putative transmembrane GTPase	1883	100
956	W74726	Homo sapiens	Human secreted protein fg949 3.	1879	100
957	Y27096	Homo sapiens	Human viral receptor protein (ACVRP).	1581	100
958	AJ222967	Homo sapiens	cystinosin	1920	100
959	Y53052	Homo sapiens	Human secreted protein clone df202 3 protein	587	100
			sequence SEQ ID NO:110.		
960	G02694	Homo sapiens	Human secreted protein, SEQ ID NO: 6775.	283	100
961	AF151855	Homo sapiens	CGI-97 protein	1214	96
962	U26592	Homo sapiens	diabetes mellitus type I autoantigen	250	65
963	AL050306	Homo sapiens	dJ475B7.2 (novel protein)	3796	100
964	AF078859	Homo sapiens	PTD004	2089	100
965	AB020315	Homo sapiens	homologue of mouse dkk-1 gene: Acc#	1466	100
		•	AF030433	1	
966	X04571	Homo sapiens	precursor polypeptide (AA -22 to 1185)	6580	99
967	AF146019	Homo sapiens	hepatocellular carcinoma antigen gene 520	993	99
968	AF071002	Homo sapiens	minK-related peptide 1; MiRP1	632	100
969	AB021227	Homo sapiens	membrane-type-5 matrix metalloproteinase	3545	100
970	AF180920	Homo sapiens	cyclin L ania-6a	1579	100
971	AF105365	Homo sapiens	K-Cl cotransporter KCC4	5621	99
972	AF083248	Homo sapiens	ribosomal protein L26 homolog	739	100
973	AJ132429	Homo sapiens	hyperpolarization-activated cyclic nucleotide	6295	100
			gated cation channel hHCN4		
974	W61619	Homo sapiens	Clone HTPEF86 of TM4SF superfamily.	454	100
975	AF155100	Homo sapiens	zinc finger protein NY-REN-21 antigen	2261	100
976	AF275948	Homo sapiens	ABCA1	11763	99
977	AB026891	Homo sapiens	cystine/glutamate transporter	2552	100
978	AF117657	Homo sapiens	thyroid hormone receptor-associated protein	3348	99
			complex component TRAP80	}	1
979	AF044201	Rattus	neural membrane protein 35; NMP35	1570	92
		norvegicus			
980	AF119297	Homo sapiens	neuroendocrine-specific protein-like protein 1	1170	99
981	AF155652	Homo sapiens	potassium channel modulatory factor	1983	99
982	W88499	Homo sapiens	Human stomach carcinoma clone HP10412-	1553	99
	<u> </u>		encoded protein.		
983	Z56281	Homo sapiens	interferon regulatory factor 3	2012	98
984	AB026125	Homo sapiens	ART-4	2160	100
985	Y14482	Homo sapiens	Fragment of human secreted protein encoded by	172	70
		<u> </u>	gene 17.		<u> </u>
986	AB023888	Homo sapiens	b-chemokine receptor CCR4	1895	100
987	W27291	Homo sapiens	Human H1075-1 secreted protein 5' end.	712	100
988	AF153450	Manduca	juvenile hormone esterase binding protein	226	32
	1	sexta		 	100
989	G03697	Homo sapiens	Human secreted protein, SEQ ID NO: 7778.	194	88
990	AF204159	Homo sapiens	potassium large conductance calcium-activated	1486	100
001	00000		channel beta 3a subunit	1	1.00
991	G02061	Homo sapiens	Human secreted protein, SEQ ID NO: 6142.	558	99
992	AL031266	Caenorhabditi	VM106R.1	327	40
000	11//2/2	s elegans	1	1000	100
993	Y66749	Homo sapiens	Membrane-bound protein PRO1124.	4730	99
994	G01246	Homo sapiens	Human secreted protein, SEQ ID NO: 5327.	141	77
995	AF133845	Homo sapiens	corin	5811	99
996	AFI17756	Homo sapiens	thyroid hormone receptor-associated protein	4999	100
			complex component TRAP150	L	
997	W62066	Homo sapiens	Human stem cell antigen 2.	284	93
998	Y87173	Homo sapiens	Human secreted protein sequence SEQ ID	725	100
	j .	 	NO:212.	1	1
	7740055		L Amino soid compans of protein DDC7762	1 1444	199
999	Y13379	Homo sapiens	Amino acid sequence of protein PRO263.	1654	
999 1000	Y13379 Y95008 AF190167	Homo sapiens Homo sapiens	Human secreted protein vf3_1, SEQ ID NO:56. membrane associated protein SLP-2	676	47

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:		ļ		Score	
1002	G01234	Homo sapiens	Human secreted protein, SEQ ID NO: 5315.	398	96
1003	W73420	Homo sapiens	Human secreted protein encoded by Gene No.	2150	100
		 	24.		
1004	X12791	Homo sapiens	19kD SRP-protein (AA 1 - 144)	742	100
1005	M23323	Homo sapiens	membrane protein	642	100
1006	X63745	Homo sapiens	KDEL receptor	326	98
1007	Y35997	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 382.	824	99
1008	AB032918	Hylobates moloch	dopamine receptor D4	92	35
1009	Y91680	Homo sapiens	Human secreted protein sequence encoded by gene 81 SEQ ID NO:353.	1372	99
1010	AL136125	Homo sapiens	dJ304B14.1 (novel protein)	825	98
1011	G03733	Homo sapiens	Human secreted protein, SEQ ID NO: 7814.	379	98
1012	Y17531	Homo sapiens	Human secreted protein clone BL205 14 protein.	818	97
1013	G00724	Homo sapiens	Human secreted protein, SEQ ID NO: 4805.	462	100
1014	AF288092	Naegleria gruberi	haem lyase	114	37
1015	AB045292	Homo sapiens	M83 protein	3867	99
1016	X15940	Homo sapiens	ribosomal protein L31 (AA 1-125)	644	100
1017	Y94873	Homo sapiens	Human protein clone HP02632.	1876	100
1018	AL024498	Homo sapiens	dJ417M14.1 (novel protein)	589	100
1019	X83425	Homo sapiens	Lutheran blood group glycoprotein	3054	99
1020	W03516	Homo sapiens	Prostaglandin DP receptor.	1864	100
1021	G03960	Homo sapiens	Human secreted protein, SEQ ID NO: 8041.	398	100
1022	Y91689	Homo sapiens	Human secreted protein sequence encoded by gene 93 SEQ ID NO:362.	768	100
1023	AE000660	Homo sapiens	hADV36S1	573	100
1024	AF132965	Homo sapiens	CGI-31 protein	1550	100
1025	W92380	Homo sapiens	Human TR-interacting protein S103a.	1466	97
1026	R66278	Homo sapiens	Therapeutic polypeptide from glioblastoma cell line.	830	100
1027	X65614	Homo sapiens	S100P calcium-binding protein	476	100
1028	Y41741	Homo sapiens	Human PRO704 protein sequence.	1323	100
1029	AJ001014	Homo sapiens	RAMPI	806	100
1030	W63682	Homo sapiens	Human secreted protein 2.	1354	99
1031	AK023007	Homo sapiens	unnamed protein product	766	100
1032	W97900	Homo sapiens	Human SR-BI class B scavenger.	2672	99
1033	Y82453	Homo sapiens	Human TGC-440 secretory protein SEQ ID NO:1.	639	99
1034	Y73473	Homo sapiens	Human secreted protein clone yd178_1 protein sequence SEQ ID NO:168.	752	93
1035	Y86468	Homo sapiens	Human gene 48-encoded protein fragment, SEQ ID NO:383.	96	90
1036	U09813	Homo sapiens	mitochondrial ATP synthase subunit 9 precursor	698	100
1037	AJ242832	Homo sapiens	calpain	3699	99
1038	X66403	Homo sapiens	acetylcholine receptor epsilon subunit CHRNE	2574	100
1039	AJ242730	Homo sapiens	polyhomeotic 2	1310	100
1040	AF169968	Mus musculus	DNA binding protein DESRT	1453	80
1041	X52563	Bos taurus	permability increasing protein	383	29
1042	G00368	Homo sapiens	Human secreted protein, SEQ ID NO: 4449.	75	50
1043	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	60	53
1044	M94582	Homo sapiens	interleukin 8 receptor B	1850	100
1045	AL080239	Homo sapiens	bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit))	1704	50
1046	AF125101	Homo sapiens	HSPC040 protein	580	100
1047	W74809	Homo sapiens	Human secreted protein encoded by gene 81 clone HMWDN32.	176	100
1048	AL022238	Homo sapiens	dJ1042K10.4 (novel protein)	2201	100
1049	W88667	Homo sapiens	Secreted protein encoded by gene 134 clone HAIBP89.	1559	99
1050	AF097518	Homo sapiens	liver-specific transporter	2820	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1051	W78324	Homo sapiens	Fragment of human secreted protein encoded by gene 81.	1318	98
1052	Y21851	Homo sapiens	Human signal peptide-contianing protein (SIGP) (clone ID 2328134).	1643	95
1053	AL163815	Arabidopsis thaliana	putative protein	661	62
1054	Y76200	Homo sapiens	Human secreted protein encoded by gene 77.	262	100
1055	AJ276567	Homo sapiens	TC10-like Rho GTPase	1160	100
1056	Y27620	Homo sapiens	Human secreted protein encoded by gene No. 54.	154	96
1057	D14530	Homo sapiens	ribosomal protein	745	100
1058	AF132000	Homo sapiens	TADA1 protein	1132	100
1059	AL031778	Homo sapiens	dJ34B21.1 (novel BZRP (benzodiazapine receptor (peripheral) (MBR, PBR, PBKS, IBP, Isoquinoline-binding protein)) LIKE protein)	920	100
1060	AF227135	Homo sapiens	candidate taste receptor T2R9	134	33
1061	Y27575	Homo sapiens	Human secreted protein encoded by gene No. 9.	1392	100
1062	Z11697	Homo sapiens	HB15	1088	100
1063	AF123757	Homo sapiens	putative transmembrane protein	819	100
1064	AF155135	Homo sapiens	novel retinal pigment epithelial cell protein	2932	99
1065	Y41674	Homo sapiens	Human channel-related molecule HCRM-2.	936 2575	100
1066	AJ250042	Homo sapiens	Rab5 GDP/GTP exchange factor homologue	770	85
1067	Y36087	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 472.	301	100
1068	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124. Human secreted protein clone mc300 1 protein	301	100
1069	Y94959	Homo sapiens	sequence SEQ ID NO:124.		99
1070	W64535	Homo sapiens	Human leukocyte cell clone HP00804 protein.	2014 148	50
1071	X03145	Homo sapiens	pot. ORF III	821	91
1072	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	249	62
1073 1074	X82200 G03213	Homo sapiens Homo sapiens	gpStaf50 Human secreted protein, SEQ ID NO: 7294.	99	47
1074	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	506	55
1075	G03187	Homo sapiens	Human secreted protein, SEQ ID NO: 7268.	424	98
1077	L25899	Homo sapiens	ribosomal protein L10	332	76
1078	Y91447	Homo sapiens	Human secreted protein sequence encoded by gene 48 SEQ ID NO:168.	898	97
1079	G01862	Homo sapiens	Human secreted protein, SEQ ID NO: 5943.	290	89
1080	AB039723	Homo sapiens	WNT receptor frizzled-3	1376	92
1081	AB020527	Homo sapiens	Na/PO4 cotransporter homolog	269	100
1082	L13802	Homo sapiens	ribosmal protein small subunit	499	80
1083	W75098	Homo sapiens	Human secreted protein encoded by gene 42 clone HSXBI25.	143	81
1084	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	83	51
1085	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	88	43
1086	AF090942	Homo sapiens	PRO0657	124	64
1087	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	129	41
1088	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	126	36
1089	AF140631	Homo sapiens	G-protein coupled receptor 14	364 114	<u>82</u> <u>32</u>
1090 1091	G04063 S72304	Homo sapiens Mus sp.	Human secreted protein, SEQ ID NO: 8144. LMW G-protein	146 .	83
1091	W88708	Homo sapiens	Secreted protein encoded by gene 175 clone HEMAM41.	405	100
1093	W85612	Homo sapiens	Secreted protein clone fh123_5.	4358	97
1093	Y53012	Homo sapiens Homo sapiens	Human secreted protein clone pm514_4 protein sequence SEQ ID NO:30.	1013	99
1095	Y92345	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-62.	409	100
1096	AF090942	Homo sapiens	PRO0657	147	60
1090	L24521	Homo sapiens	transformation-related protein	166	58
1098	X56932	Homo sapiens	23 kD highly basic protein	490	70
1099	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	83	35
1100	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	149	59
• •			clone HTDAD22.	<u> </u>	

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	A E110051	137	DD 0.1500	Score	<u></u>
1101	AF119851	Homo sapiens	PRO1722	183	72
1102	G04086	Homo sapiens	Human secreted protein, SEQ ID NO: 8167.	207	62
1103	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	91	52
1104	X74856	Mus musculus	ribosomal protein L28	128	69
1105	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	130	62
1106	G03133	Homo sapiens	Human secreted protein, SEQ ID NO: 7214.	122	48
1107	G03040	Homo sapiens	Human secreted protein, SEQ ID NO: 7121.	69	43
1108	AF039942	Homo sapiens	HCF-binding transcription factor Zhangfei	744	99
1109	AF201951	Homo sapiens	high affinity immunoglobulin epsilon receptor beta subunit	738	94
1110	AF111108	Mus musculus	transient receptor potential 2	223	79
1111	AF119900	Homo sapiens	PRO2822	144	59
1112	Y16589	Homo sapiens	A protein that interacts with presenilins.	265	39
1113	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	178	67
1114	Y02999	Homo sapiens	Fragment of human secreted protein encoded by gene 121.	164	63
1115	Y30811	Homo sapiens	Human secreted protein encoded from gene 1.	1217	99
1116	X51394	Xenopus laevis	APEG precursor protein	130	40
1117	M27826	Homo sapiens	neutral protease large subunit	442	65
1118	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	72	60
1119	G03602	Homo sapiens	Human secreted protein, SEQ ID NO: 7432. Human secreted protein, SEQ ID NO: 7683.	491	97
1120	Y35906	Homo sapiens			
1120	<u> </u>		Extended human secreted protein sequence, SEQ ID NO. 155.	244	97
	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	122	65
1122	Y00337	Homo sapiens	Human secreted protein encoded by gene 81.	110	90
1123	AF084830	Homo sapiens	two pore domain K+ channel; TASK-2	703	94
1124	AF212862	Homo sapiens	membrane interacting protein of RGS16	442	88
1125	W64469 ·	Homo sapiens	Human secreted protein from clone CW795_2.	191	53
1126	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	154	100
1127	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	165	100
1128	Y84320	Homo sapiens	Human cardiovascular system associated protein kinase-1.	815	99
1129	G02105	Homo sapiens	Human secreted protein, SEQ ID NO: 6186.	88	73
1130	Y32923	Homo sapiens	Transmembrane domain containing protein clone HP01512.	700	100
1131	Y29817	Homo sapiens	Human synapse related glycoprotein 2.	260	91
1132	Y91644	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:317.	525	96
1133	Y91449	Homo sapiens	Human secreted protein sequence encoded by gene 49 SEQ ID NO:170.	542	100
1134	AB017908	Homo sapiens	4F2 light chain	2399	93
1135	X51760	Homo sapiens	zinc finger protein (583 AA)	312	55
1136	Y99426	Homo sapiens	Human PRO1604 (UNQ785) amino acid sequence SEQ ID NO:308.	917	72
1137	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	102	50
1138	AF155106	Homo sapiens	NY-REN-36 antigen	768	91
1139	AL031055	Homo sapiens	dJ28H20.1 (novel protein similar to membrane transport proteins)	117	50
1140	AF011359	Bos taurus	regulator of G-protein signaling 7	138	96
1141	Y70018	Homo sapiens	Human Protease and associated protein-12 (PPRG-12).	623	100
1142	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	113	38
1143	AB030235	Canis familiaris	D4 dopamine receptor	89	48
1144	Y94922	Homo sapiens	Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.	539	88
1145	X99962	Homo sapiens	rab-related GTP-binding protein	398	96
1146	G03807	Homo sapiens	Human secreted protein, SEQ ID NO: 7888.	168	79
1147	G03712	Homo sapiens	Human secreted protein, SEQ ID NO: 7793.	512	85
1148	Y28279	Homo sapiens	Human G-protein coupled receptor GRIR-1.	705	76
1149	U13642	Caenorhabditi	exon 5 similar to transmembrane domain of S.	247	36

NO:	rman	% Identity
Scor		
150 G03438 Homo sapiens Human secreted protein, SEQ ID NO: 7519. 117	2	
1150 G03438 Homo sapiens Human secreted protein, SEQ ID NO: 7519. 117		
1151 G01003		62
1152 G03798 Homo sapiens Human secreted protein, SEQ ID NO: 7879. 198		80
1153 X88799		63
1154		41
1155 R74272 Homo sapiens Tumour suppressor protein, p53. 341		96
1156 Y86265 Homo sapiens Human secreted protein HUSXE77, SEQ ID 99		87
NO:180 NO:180 Section NO:180		41
1158		1
1158		98
1159 G01393 Homo sapiens Human secreted protein, SEQ ID NO: 5474. 173		42
1160 W75771		57
1161		81
1162 W67816		83
Clone HCEMU42. 230		100
1163		1.00
1164 Y87252		70
1165 W64537 Homo sapiens Human liver cell clone HP01148 protein. 338 1166 AF269286 Homo sapiens HC6 Homo sapiens HC6 Homo sapiens 1167 Y14482 Homo sapiens Fragment of human secreted protein encoded by gene 17. 1168 D90789 Escherichia Dipeptide transport system permease protein 411 1169 R63783 Homo sapiens TG0847 protein. 344 1170 Y45274 Homo sapiens Human secreted protein encoded from gene 18. 478 1171 D64154 Homo sapiens Human secreted protein encoded from gene 18. 478 1172 AB026256 Homo sapiens Organic anion transporter OATP-B 311 1173 G00357 Homo sapiens Human secreted protein, SEQ ID NO: 4438. 60 1174 D87717 Homo sapiens ribosomal protein 391 1175 M64716 Homo sapiens ribosomal protein 391 1176 R08330 Homo sapiens Human IL-7 receptor clone H6. 285 1177 L06505 Homo sapiens ribosomal protein 12 1178 A7251885 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 155 1180 G01207 Homo sapiens Human secreted protein, SEQ ID NO: 5288. 282 1181 AF181856 Rattus Rattus RNA selenocysteine associated protein 249 1182 AF161524 Homo sapiens Human secreted protein, SEQ ID NO: 7870. 282 1184 Y02671 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates dopamine receptor D4 96 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178		31
1165 W64537 Homo sapiens Human liver cell clone HP01148 protein. 338 1166		
1166		82
1167 Y14482		64
Secherichia		51
Dipeptide transport system permease protein Dipeptide transport system permease protein Dipeptide transport system permease protein Dipeptide transport system permease protein Dipeptide transport		
Coli DppC.		90
1170		İ
1170		90
1171 D64154 Homo sapiens Mr 110,000 antigen 347		98
1172		96
1173 G00357		67
1174 D87717 Homo sapiens Similar to human GTPase-activating protein(A49869) 1175 M64716 Homo sapiens ribosomal protein 391 1176 R08330 Homo sapiens Human IL-7 receptor clone H6. 285 1177 L06505 Homo sapiens ribosomal protein L12 242 1178 AJ251885 Homo sapiens Organic cation transporter (OCT2) 276 1179 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 155 1180 G01207 Homo sapiens Human secreted protein, SEQ ID NO: 5288. 282 1181 AF181856 Rattus tRNA selenocysteine associated protein 249 norvegicus 1182 AF161524 Homo sapiens HSPC176 138 1183 G03789 Homo sapiens Human secreted protein, SEQ ID NO: 7870. 282 1184 Y02671 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates dopamine receptor D4 96 concolor 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324 1190 G03361 Homo sapiens Human secreted protein, SEQ ID		52
1175 M64716 Homo sapiens ribosomal protein 391		59
1175 M64716 Homo sapiens ribosomal protein 391 1176 R08330 Homo sapiens Human IL-7 receptor clone H6. 285 1177 L06505 Homo sapiens ribosomal protein L12 242 1178 AJ251885 Homo sapiens organic cation transporter (OCT2) 276 1179 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 155 1180 G01207 Homo sapiens Human secreted protein, SEQ ID NO: 5288. 282 1181 AF181856 Rattus tRNA selenocysteine associated protein norvegicus 188 Homo sapiens HSPC176 138 1182 AF161524 Homo sapiens Human secreted protein, SEQ ID NO: 7870. 282 1184 Y02671 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1188 G00956 Homo sapiens Human secreted protein, SE		{
1176 R08330 Homo sapiens Human IL-7 receptor clone H6. 285 1177 L06505 Homo sapiens ribosomal protein L12 242 1178 AJ251885 Homo sapiens organic cation transporter (OCT2) 276 1179 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 155 1180 G01207 Homo sapiens Human secreted protein, SEQ ID NO: 5288. 282 1181 AF181856 Rattus tRNA selenocysteine associated protein 249 1182 AF161524 Homo sapiens HSPC176 138 1183 G03789 Homo sapiens Human secreted protein, SEQ ID NO: 7870. 282 1184 Y02671 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03		78
1177 L06505 Homo sapiens ribosomal protein L12 242 1178 AJ251885 Homo sapiens organic cation transporter (OCT2) 276 1179 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 155 1180 G01207 Homo sapiens Human secreted protein, SEQ ID NO: 5288. 282 1181 AF181856 Rattus tRNA selenocysteine associated protein 249 1182 AF161524 Homo sapiens HSPC176 138 1183 G03789 Homo sapiens Human secreted protein, SEQ ID NO: 7870. 282 1184 Y02671 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 107 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates concolor dopamine receptor D4 96 1188 G0956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258		67
1178 AJ251885 Homo sapiens organic cation transporter (OCT2) 276 1179 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 155 1180 G01207 Homo sapiens Human secreted protein, SEQ ID NO: 5288. 282 1181 AF181856 Rattus norvegicus tRNA selenocysteine associated protein 249 1182 AF161524 Homo sapiens HSPC176 138 1183 G03789 Homo sapiens Human secreted protein, SEQ ID NO: 7870. 282 1184 Y02671 Homo sapiens Human secreted protein encoded by gene 22 clone HMSJW18. 107 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates concolor dopamine receptor D4 96 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324<		72
1179 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 155 1180 G01207 Homo sapiens Human secreted protein, SEQ ID NO: 5288. 282 1181 AF181856 Rattus norvegicus tRNA selenocysteine associated protein 249 1182 AF161524 Homo sapiens HSPC176 138 1183 G03789 Homo sapiens Human secreted protein, SEQ ID NO: 7870. 282 1184 Y02671 Homo sapiens Human secreted protein encoded by gene 22 clone HMSJW18. 107 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates concolor dopamine receptor D4 96 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. <td< td=""><td></td><td>88</td></td<>		88
1180 G01207 Homo sapiens Human secreted protein, SEQ ID NO: 5288. 282 1181 AF181856 Rattus norvegicus tRNA selenocysteine associated protein 249 1182 AF161524 Homo sapiens HSPC176 138 1183 G03789 Homo sapiens Human secreted protein, SEQ ID NO: 7870. 282 1184 Y02671 Homo sapiens Human secreted protein encoded by gene 22 clone HMSJW18. 107 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates concolor dopamine receptor D4 96 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324		71
1181 AF181856 Rattus norvegicus tRNA selenocysteine associated protein 249 1182 AF161524 Homo sapiens HSPC176 138 1183 G03789 Homo sapiens Human secreted protein, SEQ ID NO: 7870. 282 1184 Y02671 Homo sapiens Human secreted protein encoded by gene 22 clone HMSJW18. 107 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates concolor dopamine receptor D4 96 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324		90
1182 AF161524 Homo sapiens HSPC176 138 1183 G03789 Homo sapiens Human secreted protein, SEQ ID NO: 7870. 282 1184 Y02671 Homo sapiens Human secreted protein encoded by gene 22 clone HMSJW18. 107 clone HMSJW18. 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates concolor dopamine receptor D4 96 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324		62
1183 G03789 Homo sapiens Human secreted protein, SEQ ID NO: 7870. 282 1184 Y02671 Homo sapiens Human secreted protein encoded by gene 22 clone HMSJW18. 107 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates concolor dopamine receptor D4 96 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324		Ì
1184 Y02671 Homo sapiens Human secreted protein encoded by gene 22 clone HMSJW18. 107 clone HMSJW18. 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates concolor dopamine receptor D4 96 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324		90
1184 Y02671 Homo sapiens Human secreted protein encoded by gene 22 clone HMSJW18. 1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates dopamine receptor D4 96		66
1185 G03797 Homo sapiens Human secreted protein, SEQ ID NO: 7878. 88 1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates concolor dopamine receptor D4 96 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324		71
1186 G03564 Homo sapiens Human secreted protein, SEQ ID NO: 7645. 118 1187 AB032905 Hylobates concolor dopamine receptor D4 96 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324		· .
1187 AB032905 Hylobates concolor dopamine receptor D4 96 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324		69
concolor concolor 1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324		46
1188 G00956 Homo sapiens Human secreted protein, SEQ ID NO: 5037. 292 1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324		37
1189 G03258 Homo sapiens Human secreted protein, SEQ ID NO: 7339. 178 1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324		
1190 G03361 Homo sapiens Human secreted protein, SEQ ID NO: 7442. 324		78
		79
		76
1191 AF117755 Homo sapiens thyroid hormone receptor-associated protein 187		70
complex component TRAP230		<u> </u>
1192 Y70455 Homo sapiens Human membrane channel protein-5 (MECHP- 202	•	67
5).		<u>L</u>
1193 G03052 Homo sapiens Human secreted protein, SEQ ID NO: 7133. 99		42
1194 G02607 Homo sapiens Human secreted protein, SEQ ID NO: 6688. 192		76
1195 W29661 Homo sapiens Homo sapiens CI542_2 clone secreted protein. 2001		98
1196 Y14104 Homo sapiens Human GABAB receptor 1d protein sequence. 239		69
1197 X61972 Homo sapiens macropain subunit iota 149		90
1198 G00534 Homo sapiens Human secreted protein, SEQ ID NO: 4615. 145		51
1199 Y86260 Homo sapiens Human secreted protein HELHN47, SEQ ID 1089		89
NO:175.		
1200 G02607 Homo sapiens Human secreted protein, SEQ ID NO: 6688. 154		57

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1201	G00838	Trama coniona	Thuman assessed marketin SEO FD NO. 4010		1.50
1201	M27826	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	404	50
		Homo sapiens	neutral protease large subunit	202	49
1203	Y73424	Homo sapiens	Human secreted protein clone yi4_1 protein sequence SEQ ID NO:70.	265	61
1204	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	625	98
1205	Y36203	Homo sapiens	Human secreted protein #75.	219	59
1206	U78111	Gallus gallus	AQ	205	57
1207	AF095448	Homo sapiens	putative G protein-coupled receptor	416	76
1208	AF116715	Homo sapiens	PRO2829	127	75
1209	AF099137	Homo sapiens	MaxiK channel beta 2 subunit	475	95
1210	AF205718	Homo sapiens	hepatocellular carcinoma-related putative tumor suppressor	423	79
1211	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	224	70
1212	G00719	Homo sapiens	Human secreted protein, SEQ ID NO: 4800.	117	44
1213	G01009	Homo sapiens	Human secreted protein, SEQ ID NO: 5090.	351	73
1214	AF090942	Homo sapiens	PRO0657	124	70
1215	Y14427	Homo sapiens	Human secreted protein encoded by gene 17	99	77
	1,		clone HSIEA14.	1	''
1216	G03905	Homo sapiens	Human secreted protein, SEQ ID NO: 7986.	173	57
1217	Y57897	Homo sapiens	Human transmembrane protein HTMPN-21.	1173	100
1218	J00194	Homo sapiens	hla-dr antigen alpha chain	454	78
1219	Y59709	Homo sapiens	Secreted protein 76-28-3-A12-FL1.	470	92
1220	W81576	Homo sapiens	EBV-induced G-protein coupled receptor (EBI-	725	100
1221	W96745	Homo sapiens	Dolypeptide. High affinity immunoglobulin E receptor-like	650	98
1222	Y35911	Homo sapiens	protein (IGERB).		
1223	Y00278	•	Extended human secreted protein sequence, SEQ ID NO. 160.	135	31
		Homo sapiens	Human secreted protein encoded by gene 21.	260	95
1224	AF161422	Homo sapiens	HSPC304	568	90
1225	U14970	Homo sapiens	ribosomal protein S5	202	95
1226	G01733	Homo sapiens	Human secreted protein, SEQ ID NO: 5814.	610	100
1227	AF099973	Mus musculus	schlafen2	333	56
1228	G01218	Homo sapiens	Human secreted protein, SEQ ID NO: 5299.	155	81
1229	AF217188	Mus musculus	YPIB	801	63
1230	AF176813	Homo sapiens	soluble adenylyl cyclase	275	100
1231	X98333	Homo sapiens	organic cation transporter	1704	100
1232	W74955	Homo sapiens	Human secreted protein encoded by gene 77 clone HOEAS24.	212	53
1233	Y94940	Homo sapiens	Human secreted protein clone yi62_1 protein sequence SEQ ID NO:86.	526	100
1234	U76618	Mus musculus	N-RAP	482	82
1235	AF044924	Homo sapiens	hook2 protein	380	97
1236	G01459	Homo sapiens	Human secreted protein, SEQ ID NO: 5540.	417	100
1237	AF000018	Homo sapiens	adapter protein	164	84
1238	W88633	Homo sapiens	Secreted protein encoded by gene 100 clone HE8EU04.	250	90
1239	W29660	Homo sapiens	Homo sapiens CH27_1 clone secreted protein.	697	98
1240	AF004161	Oryctolagus cuniculus	peroxisomal Ca-dependent solute carrier	154	52
1241	Y92710	Homo sapiens	Human membrane-associated protein Zsig24.	709	97
1242	Y95002	Homo sapiens	Human secreted protein vc34_1, SEQ ID NO:44.	908	88
1243	Y44905	Homo sapiens	Human potassium channel molecule ERG-LP2 partial protein.	325	100
1244	AF284422	Homo sapiens	cation-chloride cotransporter-interacting protein	511	97
	Y53629	Homo sapiens	A bone marrow secreted protein designated	1888	93
1245	!	1	I DIVISTIS.	1	1
1245	AB039371	Homo sapiens	BMS115. mitochondrial ABC transporter 3	389	97

SEQ	Accession No.	Species	Description	Smith- Waterman	% Identity
ID NO:	No.			Score	Identity
	A 70 70 500	<u> </u>	ID NO. 160.	550	90
1248	AF072509	Rattus norvegicus	glutamate receptor interacting protein 2	559	L
1249	AF247042	Homo sapiens	tandem pore domain potassium channel TRAAK	661	98
1250	B08974	Homo sapiens	Human secreted protein sequence encoded by gene 27 SEQ ID NO:131.	1087	97
1251	L15313	Caenorhabditi s elegans	putative	858	59
1252	Y29338	Homo sapiens	Human secreted protein clone it217_2 alternate reading frame protein.	278	75
1253	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	211	92
1254	G03074	Homo sapiens	Human secreted protein, SEQ ID NO: 7155.	294	83
1255	G01818	Homo sapiens	Human secreted protein, SEQ ID NO: 5899.	253	91
1256	AF286368	Homo sapiens	eppin-l	222	54
1257	AF220264	Homo sapiens	MOST-1	87	93
1258	G02227	Homo sapiens	Human secreted protein, SEQ ID NO: 6308.	281	78
1259	Y07970	Homo sapiens	Human secreted protein fragment #2 encoded from gene 26.	81	94
1260	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain ligand (clone 3TW).	986	100
1261	AF140674	Homo sapiens	zinc metalloprotease ADAMTS6	172	36
1262	U28369	Homo sapiens	semaphorin V	237	67
1263	Y07049	Homo sapiens	Renal cancer associated antigen precursor sequence.	288	71
1264	Y36153	Homo sapiens	Human secreted protein #25.	187	80
1265	Y78114	Homo sapiens	Human cytokine signal regulator CKSR-2 SEQ ID NO:2.	723	93
1266	Y13397	Homo sapiens	Amino acid sequence of protein PRO334.	191	100
1267	AF030558	Rattus norvegicus	phosphatidylinositol 5-phosphate 4-kinase gamma	859	95
1268	U73167	Homo sapiens	candidate tumor suppressor gene LUCA-1	159	96
1269	AF190664	Mus musculus	LMBR2	552	76
1270	AL050332	Homo sapiens	dJ570F3.1 (homolog of the rat synaptic ras GTPase-activating protein p135 SynGAP)	820	98
1271	G02126	Homo sapiens	Human secreted protein, SEQ ID NO: 6207.	131	95
1272	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	253	92
1273	AL035661	Homo sapiens	dJ568CI1.3 (novel AMP-binding enzyme similar to acetyl-coenzyme A synthethase (acetate-coA ligase))	1280	100
1274	AF064748	Mus musculus	S3-12	3523	61
1275	D17554	Homo sapiens	TAXREB107	377	78
1276	Y30715	Homo sapiens	Amino acid sequence of a human secreted protein.	643	90
1277	AF146760	Homo sapiens	septin 2-like cell division control protein	707	100
1278	Y05069	Homo sapiens	Human PIGR-2 protein sequence.	281	46
1279	X59668	Oryctolagus cuniculus	aorta CNG channel (rACNG)	267	85
1280	G01051	Homo sapiens	Human secreted protein, SEQ ID NO: 5132.	489	98
1281	G03411	Homo sapiens	Human secreted protein, SEQ ID NO: 7492.	120	43
1282	AF055084	Homo sapiens	very large G-protein coupled receptor-1	1635	100
1000	AF117814	Mus	odd-skipped related 1 protein	357	98
_		musculus	<u></u>		
1283 1284	U87318	Xenopus laevis	NaDC-2	535	60
_	U87318 AF061346	Xenopus	NaDC-2 Edp1 protein	535 452	68
1284		Xenopus laevis Mus	Edp1 protein contains transmembrane (TM) region	452	68
1284 1285	AF061346 AB030182 A13595	Xenopus laevis Mus musculus Mus musculus synthetic construct	Edp1 protein contains transmembrane (TM) region immunosuppresive protein PP15	452	68
1284 1285 1286	AF061346 AB030182	Xenopus laevis Mus musculus Mus musculus synthetic	Edp1 protein contains transmembrane (TM) region	452	68

SEQ ID NO:	No.		Description	Smith- Waterman Score	% Identity
1290	AF038563	Homo sapiens	membrane associated guanylate kinase 2	523	100
1291	AF034837	Homo sapiens	double-stranded RNA specific adenosine deaminase	468	100
1292	M15888	Bos taurus	endozepine-related protein precursor	937	87
1293	AB010692	Arabidopsis thaliana	ATP-dependent RNA helicase-like protein	636	45
1294	AF209923	Homo sapiens	orphan G-protein coupled receptor	1570	100
1295	W67828	Homo sapiens	Human secreted protein encoded by gene 22 clone HFEAF41.	504	98
1296	AC004832	Homo sapiens	similar to 45 kDa secretory protein; similar to CAA10644.1 (PID:g4164418)	648	65
1297	X80035	Oryctolagus cuniculus	cysteine rich hair keratin associated protein	575	70
1298	G02645	Homo sapiens	Human secreted protein, SEQ ID NO: 6726.	223	97
1299	Y59440	Homo sapiens	Human delta3 fragment #4.	122	32
1300	W70504	Homo sapiens	Leukocyte seven times membrane-penetrating type receptor protein JEG18.	459	81
1301	Y67315	Homo sapiens	Human secreted protein BL89_13 amino acid sequence.	3916	99
1302	M77693	Homo sapiens	spermidine/spermine N1-acetyltransferase	174	96
1303	G01331	Homo sapiens	Human secreted protein, SEQ ID NO: 5412.	254	69
1304	G01491	Homo sapiens	Human secreted protein, SEQ ID NO: 5572.	747	99
1305	AF148509	Homo sapiens	alpha 1,2-mannosidase	602	98
1306	G01658	Homo sapiens	Human secreted protein, SEQ ID NO: 5739.	333	98
1307	Y90899	Homo sapiens	D1-like dopamine receptor activity modifying protein SEQ ID NO:1.	332	98
1308	AF033120	Homo sapiens	p53 regulated PA26-T2 nuclear protein	348	52
1309	Y73388	Homo sapiens	HTRM clone 3376404 protein sequence.	147	66
1310	AF063243	Bos taurus	ribosomal protein L30	296	90
1311	AF224494	Mus musculus	arsenite inducible RNA associated protein	688	70
1312	Y73342	Homo sapiens	HTRM clone 2709055 protein sequence.	1154	100
1313	Y99419	Homo sapiens	Human PRO1780 (UNQ842) amino acid sequence SEQ ID NO:282.	1145	78
1314	AF116667	Homo sapiens	PRO1777	433	97
1315	W75100	Homo sapiens	Human secreted protein encoded by gene 44 clone HE8CJ26.	807	97
1316	AJ272078	Homo sapiens	APOBEC-1 stimulating protein	789	100
1317	AB041533	Homo sapiens	sperm antigen	2607	98
1318	U19617	Mus musculus	Elf-1	806	92
1319	U82598	Escherichia coli	ferric enterobactin transport protein	768	100
1320	D90892	Escherichia coli	SORBITOL-6-PHOSPHATE 2- DEHYDROGENASE (EC 1.1.1.140) (GLUCITOL-6- PHOSPHATE DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE).	709	100
1321	W67847	Homo sapiens	Human secreted protein encoded by gene 41 clone HPBCJ74.	601	92
1322	AJ276101	Homo sapiens	GPRC5B protein	466	93
1323	AJ276101	Homo sapiens	GPRC5B protein	504	97
1324 1325	Y58628	Homo sapiens	Protein regulating gene expression PRGE-21.	1584	100
	U91561	Rattus norvegicus	pyridoxine 5'-phosphate oxidase	1277	89
1326	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	1606	100
1327	Y32206	Homo sapiens	Human receptor molecule (REC) encoded by Incyte clone 2825826.	1531	90
1328	AF151048	Homo sapiens	HSPC214	657	85
329	Y10530	Homo sapiens	olfactory receptor	1645	100
1330	AF180681	Homo sapiens	guanine nucleotide exchange factor	4314	99
1331	AF111856	Homo sapiens	sodium dependent phosphate transporter isoform NaPi-3b	3591	99
1332	Y13583	Homo sapiens	G-protein coupled receptor	2171	100
1333	AF078866	Homo sapiens	SURF-4	1395	100

SEQ	Accession	Species	Description	Smith-	7%
ID	No.	\		Waterman	Identity
NO:				Score	
1334	Y25755	Homo sapiens	Human secreted protein encoded from gene 45.	1380	96
1335	AF152325	Homo sapiens	protocadherin gamma A5	4742	99
1336	X74070	Homo sapiens	transcription factor BTF3	639	81
1337	AF095927	Rattus norvegicus	protein phosphatase 2C	1931	95
1338	G03877	Homo sapiens	Human secreted protein, SEQ ID NO: 7958.	621	100
1339	AL008582	Homo sapiens	bK223H9.2 (ortholog of A. thaliana F23F1.8)	626	100
1340	X61615	Homo sapiens	leukemia inhibitory factor receptor	5820	99
1341	Y01519	Homo sapiens	A carcinogenesis-inhibiting protein.	7528	97
1342	AF207600	Homo sapiens	ethanolamine kinase	2372	100
1343	U54807	Rattus norvegicus	GTP-binding protein	1167	97
1344	AC020579	Arabidopsis thaliana	putative phosphoribosylformylglycinamidine synthase; 25509-29950	3283	51
1345	Y28576	Homo sapiens	Secreted peptide clone pe503_1.	944	100
1346	W74787	Homo sapiens	Human secreted protein encoded by gene 58 clone HHFHN61.	1171	100
1347	M55542	Homo sapiens	guanylate binding protein isoform I	2636	87
1348	AF183428	Homo sapiens	28.4 kDa protein	1329	100
1349	U70669	Homo sapiens	Fas-ligand associated factor 3	167	24
1350	AF295530	Homo sapiens	cardiac voltage gated potassium channel modulatory subunit	562	99

TABLE 3

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino, acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	цепсе	1	C7/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ł		 	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	į.	ì		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	i .			peptide	· .	/=possible nucleotide deletion, \=possible
	!	ł	i	sequence		nucleotide insertion
1	1351	A	2	337	1	TPSLIHQAPTPCPAGLWG/PPNGHYHGS*PGC
į	l	1	i .	Į.		HWPQAPHRA***GLLPPRWLGHGLPGGPAAP
İ	1			•		WAASQWVDGVAGRLPGPAWSWHASGAAPA
	}	1	į	Ì	}	QPGPL*LLVPGSSGLPDPRDP
2	1352	A	27	100	366	IRNSSIRPMKERETKLSAKHMITCSASYDIRGL
-			ŀ		ł	QIETT\YHHTPIRMAKIQKT/GHHQC**ECGAT
		ļ		1		GTLIHGWWGCKVVEPLGKTVWQIPK
3	1353	A	40	3	314	HASAHASVVLKDNSELEQQLGATGAYRARA
-	1					LELEAEVAEMRQMLQLEHPFVNGADKLRPD
ŀ	ĺ	1		[[SMYVHLNEL*QSLVENMLLTVVDTH\RTPI*R
ļ						SCNYTLALILFL
4	1354	A	74	2	292	TASALFSCPDGGSLAGFAGRRASFHLECLKR
1		1				QKDRGGDISQKTVLPLHLVHHQVAHTFGQAT
Ī	1		[· ·		VTCQQARQSPG*RTNPE/ALQWVLPVSDGWH
1						VLPLP
5	1355	A	78	114	850	ENCRVASNLPGVFFSEDTAQSGSYMRISAHPP
1						NAGGEVSNGPKRKLTLMLNFSLPSSGLNAGA
			ľ		1	FYALSTLLNRMVIWHYPGEEVNAGRIGLTIVI
İ	· ·		1		ļ	AGMLGAVISGIWLDRSKTYKETTLVVYIMDT
l l	İ					GGAWWCYTFYLGTGDTCG*CFITAG\TMGFF
]		1	ì		MTGYLPLGFEFAVEL\SYPESEGISSGLLNISA
] .	1	ļ		QVFGIIFTISQGQIIDNYGTKPGNIFLCVFLTLG
1]		l]	AALTAFIKADLRRQKANKETLEN
6	1356	A	81	97	376	EWFSYMLGSNMSVYHSP*SLEPLCKVLSES*A
1			1	1		YLRVPFIRILLNAR*IRKAYKRMSLEIKLLI/RE
1	1	1	1	1	ţ	*CLFQEMGLSLQWLYSARGDFFRATSRL
7	1357	A	93	2	872	TLSSACLIGDAWKELTIVAGAVSNQLLVWYP
		1		1		ATALADNKPVAPDRRISGHVGIIFSMSYLESK
1	1	1	1			GLLATASEDRSVRIWKGGDLRVPGGRVQNIG
	1	ł	i	}	1	HCFGHSARVWQVKLLENYLISAGEDCVCLV
l————				<u> </u>		

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion WSHEGEILQAFRGHQGRGIRAIAAHERQAWV ITGGDDSGIRLWHLVGRGYRGLG/DLGSLLQ VP**ARYTQGCDSGWLLATAGSD*YRGPVSL
8	1358	A	106	3	350	*RRGQVLGAAARG*TFPVLLPAGGSSWSRGL RIVCYGQWGRSCQGCPHQHSNCCCGPDPVS WEGAQLELGPAWL FSSLLSGRISTLRDETGAILIDGDPAACAPIIKF LLTEELHLRGVSIYVLRHEAQIYGITPL\VCAL
					, 	LI/CRRL*SDSCMRAALNDRGLYQVLILDGLV QCLGFVDSDSRKMVSTLT
9	1359	Α	115	49	186	QAWAIFKGKYKEGDTGGPAVWKTRLRCALN KSSEFNEGPERERMDV
10	1360	A	123	2	1249	KGCRTQEKVDRTEVIRTCINPVYSKLFTVDFY FEEVQRLRFEVHDISSNHNGLKEADFLGGME CTLGQIVSQRKLSKSLLKHGNTAGKSSITVIA EELSGNDDYVELAFNARKLDDKDFFSKSDPF LEIFRMNDDATQQLVHRTEVVMNNLSPAWK SFKVSVNSLCSGDPDRRLKCIVWDWDSNGK HDFIGEFTSTFKEMRGAMEGKQVQWECINPK YKAKKKNYKNSGTVILNLCKIHKMHSFLDYI MGGCQIQFTVAIDFTASNGDPRNSCSLHYHHP YQPNEYLKALVAVGEICQDYDSDKMFPAFGF GARIPPEYTDSHDFAINFNEDNPECAGIQGVV EAYQSCFVKAPTFTGPTNICPHSSRKVAKFRR SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP DNPGGHFV
11	1361	A	147	614	9	ACARKQLLGRTVFIWFVGQLLGGELKGYSKT NTTSSRPASSRG\TLSSSSSSSSSLTKDALPSSL KSDSTTITSGLVFPFRSLCVNPAKSSVSESVSSI KILLSSSVKYLE*KRTSCCFPDSSESKLSQLSS DERVSMGTSSRKPTNSSSSLGALKMSATS*G SGSESPTPFFLTGLQSPPSTRPREPGLTTARNS TTLTRDC
12	1362	A	177	12	416	LIPSEPALDSLVDPRVRSRKQPFVIYPVYDTAI DTKIHFSLLDGNVGEPDMSAGFCPNHKAAM VLFLDRVYGIEVQDFLLHLLEGGFLPDLRAA ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI NLLTMNVY
13	1363	A	249	535	105	WTFHRHLSPAPLIVCDQGTCVVSYYPQNIVQ MPDTQMEQGLN/HLFLDGNA*PHSVECYCPS TFEIAIKITSFVLYFHRYRAPEVLLRSSVYSSPI DVWAVGSIMAELYMLRPLFPGTSEVDEIFKIC QVLGTPKKVSTLVPKLL
14	1364	A	254	572	201	YLLTXIGNLMMLLVINADSCLRTXM*FFLGH FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT *SFFFIFSWGTNGCLLSAITYACYAAICHPLLS TMVMNRPLCTATVNATNKMGFLNSQVN
15	1365	A	257	425	68	THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM AIEFLLECDQNITKLICENT*KNIAKNI*KRRV TFTPIET*HPVKQMIKWQ*LTAWLRNRGYKKI KQTPNSETAPSVCRNLVFDKCG
16	1366	A	263	104	481	FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV
17	1367	A	298	68	208	RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM
18	1368	A	300	904	1	LVVGITGTRHHARVIFIFLVETGFPHVGQAGL ELLTSGDPPALASQSAGITGMSHCARPKGHFG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning marleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IHLK*MFYTMSQKMP*PTINLILLLIIPGNLNIF KPNMGWLGPKTAFV*KDEVLSGIPFAKGRCR WK*DY*C/LQEVTDPIMEKGKKKKRTASFFK GQPHQSTNALLRRCVR*RYHLS\TVETAGLP*KNTGHIPGQPFLFKLVFKC*NVICI**QYKW*Q NIGVKNKSFCPH*SSSPSL*FIGHHSRNF/CSFK TEPHSVVQAGGQWRNLSSLQAPPPGLMPLSR ISLMSSWDYRRPPQ
19	1369	A	302	3	445	NSPSRWAKIQMFEHTFCG*GCG/ER/NVHIHCS WICRLRPLLWRAVREYLSKLKNAELSFDPGV SLLRIYAIDMPTSI*DEKEALLFAFLAFHE*HC KSRIWAVIQ/CIHLWDWLRKL*CFHRMKFYA AV*NKPRHILLSHIWKDVQNILLK
20	1370	A	304		1339	FFFCGKEVPLFEQNKHPGPRATTSPGA/HARA LLSAGEFTAGVGLSP*AIHSFVWLCTFIQHGA GGPCHQPGGSPGPWMHTTQAGHLWEGAYPG GSSTWHQVPGQLGGSWGPRERSLLGSFIKCSP CPHPPGFRLWMSPNQKPPTENPGVMGRVWR LMPGESPLIWEAEGKEDHLSPEGQGHSE/PVA PLHSSLGNTVKP*PKNQKPKQNRSRHGQ\GF MAGQGQSRPAAR*PPCPALTPASHSAGTWPP RICRTVPGGPCPSPSGFRSCRR*GFSA*TRSWP DAEPPSTPDTAPRCCTQSDTSSQGPQ*S*WRR CRALPGRLCSAPAAGLRRARPRLSESRRGNSP PASPAAASARCPSWGPSCPARPPSRPAAGTEP AAPSRCTAWLRGEREPGPRPPGRRPRSGRGP VSFAPEVLSLPAVRQTKSWRWRNEEEITRPW ALVRSRGG
21	1371	A	326	799	1587	GSQVLPPPPSQDSATLPQDA*GPRAAPGQPVC E*GLQGAGVRRLRGEVLCQPQP*GAL*EQCLP HLSFSPRQGAAPDTEPSAWGPAPTGATGPGLP LRHVRLFSAGAPRGAATPCPPALLHGPAWPP ARPMFRGHPPVRPLGPWGKVAAGPRALCLA GVPAVQGECATKPSG*GL*PAHLRGPPGPEVL QWHWQLSAGRDPVPAEDPPL*EGPLGPGGPA AAQAEPGADPEPEDKDQAAESRPAGAMSLSA QGSGPVGGQGLR
22	1372	A	327	146	652	PHLENPHPEHSFPGAPLT*STLSWSILSPREPSP GAPCYPGHPHLENPHLEHLTWRTVTWSTLL PGAPCYPEHPHLEHPLTWSTPHLEHPSPGEPL SCRTPTRSILHRDHPLP*CLSTEESPI*GWGSLP APPSTPLVLDVAPPGPQPASSCPGRDSCYSVP GTVVSP
23	1373	A	348	397	2	CIVSSCQGTRKPCHLEDANKINKQSPTLEKIES LQESL*VKQ*LIVAEKYVQILHPRKKYFQRPL NNEKRKMKKRKEEKKKCRERMQRRSKWRR EEKKE*RREE\EERKKEKEDRKERRKETSPRG SRRLLRD
24	1374	A	362	170	352	GRALDTAAGSPVQTAHGLPSDALAPLDDSMP WEGRTTAQWSLHRKRHLARTLLVSRVRGPQ
25	1375	A	384	373	128	YLITTILETGYLWKNRHSDQ*KRTENPERDQH KYPKVDFCKSNSMKNRLCNKWHWTNWIFTD KKINLNLKPHTKLTPNIKKN
26	1376	A	397	383	165	EVKNTNPFIFSGTNLTIWIRSI*RKSDEINQRTK *MEKYSISLDRRLNTVKMSFLPNLIYKFNTISI KIPANF
27	1377	A	406	103	380	KSKATGYMVNI*KLIV\FLYANDEQLEIEMNK IVP\FNGSKNKIAFTNLTKYQNIQNRHAENYKI LVNKIEDLNKWRNVLLSWIGRRNIINTMT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
28	1378	A	408	14	427	TICTNKFNNLDEIK/FLERHKLSKLTQEEVENL ITLKTSRETELVINK*VIPHKEKPGPDSFTGEF YQTFKEEL/II/ILHKLFQTIKYGRILPNSVYETSI TLKPKPEKDL\KENYRPLPLSNIDAK\LNKTLA NRI**HIR
29	1379	A	434	395	128	IYSKMCMERQRLNN*ILKKNKVRGIAVPDVK VYYKPTVIK/TSWIL*KDSHIVEWNRLENLEID PN/IKRLILDKGAEATEWRKDSFFRQWQ
30	1380	A	455	2	228	FFFETESHSVTQAGVQWCNPGFKRFSCFGLSS SWDYRYAPPRP\ANF*FLVETGFYYVAQAGL KLLSPGDLPALAS
31	1381	A	462	393	2	QLMFDKGVKNIH\WGWTPPFTK*YWKNWISI CRRMNLNPYLSRYIKINSR\KDLTVRPEPIKLV EENTGKTIQDTGLGK*FIAKTSKAQSTKTNK* KRQTRYIKLK\KKSTASKENNRVKRQPLE*EK IFAN
32	1382	A	474	125	471	VKPYEIAVFLVKPIEYK*HILSDPAIPLSGI*LK EIKAYT/RRICTPMFAAPVSVIA/RN*KQSK/CQ KQ*YVHRMEYYTTIKRSEILICTTTWVDFRNT ILRETDRIHKTTYDVISLI
33	1383	A	488	1825	2	KSACSFICSEEQPASPSPLKPGTYASETNPPDP HAAGPRRDSSEAETRRPRGA/DGSGTVVKGT PGSPAPPCSWGHGG\ETEGAG*CPAAPGTDLR APGGSAGS*\GLPSAGGSRGKKGWRAAGRQP STR*GRPGRHGGRGE*AGHPEPRQSALQSAG L/ASSPEPMGAALAEDGSGDSRGAGPRPQE*P PSVLSRS\GS*G*G*AASGTASSPRSHSSRLGPP SAGFHGLRCGQPPFAAAPPGPWPGTGRPAGG AGSPPAAAGTAPPATRGAQSRRQNRTAGRNA SPQTAAGAGSPVQWALSRATG*TGETGSWC AGGTHQATHLTAAWVCPPTWSVRPGGSGPA AGLGR*GRHPAQSPPLPVPRG*PAWPQEAPSP SPASSEVALSSGSCWPDQAPGPARGSPPAPLA PAWPAAGRGRQR*GRQSAHPPPRR*STAVSL SGTS*WRRSP*AGTRTQQC*SPWLVPACSSRP L*RGTRRPSTQQSPQTTGTTGRSAGPGHPRS* GGRSPAGTGHLGAQTVASPH*GHWPTALSCL WASASPPGPEAPPQTGACIGTNCRYRAASAR RSSVAPACA*GWQ*AGSPPAVLRGPP*RVRER GALTHRPRAPDE
34	1384	A	497	422	2	APGASVGRAQAAEG*RGGPTGRPPSALGVS/E AGRAGRAGEGRPVPPAYPLCKSAQTSGPPKA RLS\PPLASCGGRGPPGGAACATCAPPAGPAR SSRCRRRSPPE*GPR*PSRPARPSPGSAASRRQ KLTPCRCQFRGLCA
35	1385	A	509	156	475	PTPYPGE*QAAFLLRGPGLRPPA/DPSLR/HRN LTELVVAVTDENIVGLFAALLAERRVLLTAS KLSTLTSCDHAFCALLYPMRWEHVLIPTLPPH LLDYC*CPPLPRT
36	1386	A	512	3	1631	FFFSFVCHLYCVSPTPGPHGRLATWL/PGLLA FLGLAAGGQTLCPAGELPGHARAQASGAPGS VLIAVPGRRRVHTCGPGPAAPSTRGECPPPAL GHTRPARPRPV\PFAPAVPQEPGGQGHGAA/P PATGHSAPRGCPPARAAPTGSATPAPPPAACA AFHSAWSVPPAGRQQG*RVPAPAFRRTTPGT PGQHLLDRPGAPPAQGSGPAPAPPPRLAGPA GPAAPPPGPPAASWHSSLSKSSSSL\GWSPPLP VGPGSLQ*TPPPQGPHLSGSCGGTSSWRGQR AAVARRLRSWNACGLSRVAGRSSASYPGRE

SEQ ID Not of Not of nucleotide coide sequence Not of nucleotide coide sequence Not of nucleotide coide sequence Not of nucleotide USSN Location Loca	,
nucle- cotide seq- uence	
cottide sequence USSN location 09/496 corresponding to last amino acid residue form fo	
Sequence 99/496 correspondi olast amino acid residue of peptide sequence correspondi sequence correspondi correspondi sequence correspondi cor	
uence 914 ng to first amino acid residue of peptide residue of peptide residue of peptide sequence peptide sequence sequ	
amino acid residue of peptide sequence T=Threonine, V=Valine, W=Trytophan, v=Stop codon, peptide Sequence T=Threonine, V=Unknown, *=Stop codon, peptide Sequence Sequence Sequence Typosible nucleotide deletion, t=possible nucleotide insertion GRFSOSQ*PAGPFGMRGCCLRGW*PSSS GPGPHPASTWLRAGKTGPSPPACGCA*L VSAAPQSPRTRCPRGCAAAAGLCVLAA HGAIGLGVRVHTIORVHIH*GAG/GCQTI LRSLPVLGLPAPRCPVSAHPWHRRSGSS ARLVPRHPAPGCP**TG*PLITGFPEP*A* NHQAVGLEASGALQAGHRDELPTMVQI SPDVPLKGRPHAP	
peptide	
Sequence	
GRPSQSQ*PAGPPGMRGCCLRGW*PSSS GPGPHPASTWLRAGKTGPSPPAGGCAT	
GPGPHPASTWLRAGKTGPSPPAGGCA1L	_
VSAAPQSPRTRCPRGCAAAAGLCVLAA	
HGA\GL\PGVRVHT\QRVHIH*GAG\/GC\PTI LRSLP\VLGLPAPRC\PV\SAHP\WIRRSGS\S\RIV\PRI\PA\GL\PAPRC\PV\SAHP\WIRRS\GS\S\RIV\PRI\PA\GL\	
LRSLPVLGLPAPRCPVSAHPWHRRSGSS ARLVPRHPAPGCF**TG*PULITGFPEF*A* NHQAVGLEASGALQAGHRDELPTMVQI SPDVPLKGRPHAP SPDVPLKGRPHAP SPDVPLKGRPHAP SPDVPLKGRPHAP SPDVPLKGRPHAP RLNSTHLYACGTHAFQPLCAAIDAEAFT FEEGKEKCPYDPARGFTGLIIDGGLYTATA FEEGKEKCPYDPARGFTGLIIDGGLYTATA FEEGKEKCPYDPARGFTGLIIDGGLYTATA FEEGKEKCPYDPARGFTGLIIDGGLYTATA FEEGKEKCPYDPARGFTGLIIDGGLYTATA FEEGKEKCPYDPARGFTGLIIDGGLYTATA FEEGKEGKEPGVIGRESPACLKSLSN VAPEFVFSVLVRESKASAVGDDDKVYYF RATEKESGSFTQSRSSHRVARGIPPL ATTEKESGSFTQSRSSHRVARGIPPL KDKKEVGFQSIQALMQTC\GEKVMAD QDLPRFLQLLCEGHNNDFQNYLRTQTGI NIIICTVDYLLRLQESI SPD	
ARL VPRHPAPGCP*TG*PLITGFPEP*A* NHQAVGLEASGALQAGHRDELPTMVQI SPDYPLKGRPHAP 37 1387 A 620 828 1 FRLPLAAGA/RGAAEPRVAVSMAPDPSA WEASPEMQSKCHQKGKNNQTECFNHVI RLNSTHLYACGTHAFQPLCAAIDAEAFT FEEKEKEKCPYDPARGFTGLIIDGGLYTAT FRSIPDIRRSRHPHSLRTEETPMHWLNG* AQDDGG*GTISSFLLPWPADHPTPKSPGI SIPVCQV/RQQPQSGGKESPACLKSLSN VDAEFVFSVLVRESKASAVGDDDKVYYF RATEKESGSFTQSRSHRVARGIPPL 38 1388 A 739 1 427 FRAMVSTLKLGISILNGGNAEVQ/QGN TSEEGKEG*EVPV*LPVSPPLPRLQKML KDKKEVGFFQSIQALMQTC\GEKVMADI QDLPFLQLLCEGHNNDFQNYLRTQTG NIIICTVDYLLRLQESI 39 1389 A 767 1 1030 TLDLTGPLLLGGVPNVPKDFRGRNRQFC RNLSVDGKNVDMAGFIANNGTREGCA# FCDGRRRQNGGTCVNRWNMYLCECPL KNCEQGEWPASSIPPVTAAWEALLLDVI VRGLHIQVRQPLVYYAAFTVDSHRPLQI RRAPAPASGVPSPSGVGWNR*AGPAEPS ATVIISVPWYLGLMFRTRKEDSVLMEA PTSFRLQVTGAPCHQGTC*VGARGRDPM LRVTDGEWHHLLLEKNVKEDSEMKHL TLDYGMDQVSWHLHILLWG*TLPPAQG SEDKVSVRRGFRGCMQVRGGCGGRGE QAAPRL 40 1390 A 801 69 399 HKIIIIHKEDLNKWKYILCSGMERLSTVM PQIIYKFNA*Q\VILKFTW*E*GAKITILRF RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	
NHQAVGLEASGALQAGHRDELPTMVQI SPDYPLKGRPHAP 37	
SPDYPLKGRPHAP	
1387 A 620 828 I FRLPLAAGA/RGAAEPRVAVSMAPDPSA	LLDH
WEASPEMQSKCHQKGKNNQTECFNHVI RINSTHLYAGTHAFQPLCAADDAEAFT FEEGKEKCPYDPARGFTGLIDGGLYTAT FEEGKEKCPYDPARGFTGLIDGGLYTAT FESIPDIRRSRHPHSLRTEETFMHWLING* AQDDGG*GTISSFLLPWPADHPTPKSPGI SIPVCCQVRGQPQSGGKESPACLKSLSNK UDAEFVFSVLVRESKASAVGDDDKYYYF RATEKESGSFTQSRSSHRVARGIPPL RATEKESGSFTQSRSSHRVARGIPPL SEEGKEG*EVPV*LPVSPPLPRPLQKML KDKKEVGFFQSIQALMQTCGEKVMADD QDLFRFLQLLCEGHNNDFQNYLRTQTGI NIIIICTVDYLLRLQESI NIIICTVDYLLRLQESI NIIICTVDYLLRLQESI NIIICTVDYLLRLQESI KNCEQGEWPASIPPVTAAWEALLLDVI VRGLHQVRQPLVYYAAFTVDSHRPLQF RRAPAPASGVPSPSGVGWDR*AGPAEPS ATVIISVPWYLGLMFRTRIKEDSVLMEA PTSFRLQVTGAPCHQGTC*VGARGRDPN LRVTTDGEWHHLLIELKNVKEDSEMKHL TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGE QAAPRL WIKINIKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*QIVILKFTW*E*GAKITILRI RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM A	
RLNSTHLYACGTHAFQPLCAAIDAEAFT FEEGKEKCPYDPARGFTGLIIDGGLYTAA FRSIPDIRRSRHPHSLRTEETPMHWLNG* AQDDGG*GTISSFLLPWPADHPTPKSPGI SIPVCCQVRGQPQSGGKESPACLKSLSNG VDAEFVFSVLVRESKASAVGDDDKVYYF RATEKESGSFTQSRSSHRVARGIPPL 38 1388 A 739 1 427 FRAMVSSTLKLGISILNGGNAEVQ/QGNI TSEEGKEG*EVPV*LPVSPPLPRPLQKML KDKKEVGFFQSIQALMQTCCGEKVMADI QDLPRFLQLLCEGHNNDFQNYLRTQTGI NINICTVDYLLRLQESI 39 1389 A 767 I 1030 TLDLTGPLLLGGVPNVPKDFRGRNRQFG RNLSVDGKNVDMAGFIANNGTREGGAA FCDGRRRQNGGTCVNRWNMYLCECPLI KNCEQGEWPASSIPPVTAAWEALLLDVI VRGLHIQVRQPLVVYAAFTVDSHRPLQI RRAPAASGVPSPSGVGWDR*AGPAEPS ATVIISVPWYLGLMFRTRKEDSVLMEA* PTSFRLQVTGAPCHQGTC*VGARGRDPN LRVTDGEWHHLLIELKNVKEDSEMKHL TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGEA QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*QVUILKFTW*E*GAKITILRF RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	
FEEGKEKCPYDPARGFTGLIIDGGLYTAT FRSIPDIRRSRHPHSLRTEETIMHWLNG* AQDDGG*GTISSFLLPWPADHPTPKSPGI SIPVCCQVRGQPQSGKESPACLKSLSNN VDAEFVFSVLVRESKASAVGDDKVYYF RATEKESGSFTQSRSSHRVARGIPPL 38 1388 A 739 1 427 FRAMVSSTLKLGISLINGGNAEVQ/QGNI TSEEGKEG*EVPV*LPVSPPLPRPLQKML KDKKEVGFFQSIQALMQTCCGEKVMADI QDLFFFLQLLCEGHNNDFQNYLRTQTGI INIIICTVDYLLRLQESI 39 1389 A 767 1 1030 TLDLTGPLLLGSVPNVPKDFRGRNRQFC RNLSVDGKNVDMAGFIANNGTREGCAA FCDGRRRQNGGTCVNRWNMYLCECPLI KNCEQGEWPASSIPPVTAAWEALLLDVI VRGLHIQVRQPLVVYAAFTVDSHRPLQI RRAPAPASGVPSPSGVGWDR*AGPAEPS ATVIISVPWYLGLMFRTRKEDSVLMEA' PTSFRLQVTGAPCHQGTC*VGARGRDPN LRVTDGEWHHLLIELKNVKEDSEMKHL TLDYGMDQVSWHLHLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGEA QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*QVVILKFTW*E*GAKITILRF RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	
FRSIPDIRRSRHPHSLRTEETPMHWLNG*	
AQDDGG*GTISSFLLPWPADHPTPKSPGF SIPVCCQVRGQPQSGKESPACLKSLSNN DAEFFSVLVRESKASAVGDDDKVYFI RATEKESGSFTQSRSSHRVARGIPPL RATEKESGSFTQSRSSHRVARGIPPL RAMVSSTLKLGISLINGGNAEVQ/QGNF TSEEGKEG*EVPV*LPVSPPLPRPLQKML KDKKEVGFFQSIQALMQTCCGEKVMAD QDLFRFLQLLCEGHNNDFQNYLRTQTG NIIICTVDYLLRLQESI SIPVE A 767 1 1030 TLDLTGPLLLGGVPNVPKDFRGRNRQFC RNLSVDGKNVDMAGFIANNGTREGCAA FCDGRRRQNGGTCVNRWNMYLCECPLI KNCEQGEWPASSIPPVTAAWEALLLDV VRGLHIQVRQPLVVYAAFTVDSHRPLQF RRAPAPASGVPSPSGVGWDR*AGPAEPS ATVIISVPWYLGLMFRTRKEDSVLMEA PTSFRLQVTGAPCHQGTC*VGARGRDPP LRVTDGEWHHLLEKNVKEDSEMKH TLDYGMDQVSWHLHLLWG*TLPPAQG SEDKVSVRRGFRGCMQVRGGCGGRGE QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVM PQITYKFNA*QIVILKFTW*E*GAKITILRF RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	
SIPVCCQVRGQPQSGKESPACLKSLSN(
NDAEFVFSVLVRESKASAVGDDDKVYYYE RATEKESGSFTQSRSSHRVARGIPPL	
RATEKESGSFTQSRSSHRVARGIPPL	
1388	
TSEEGKEG*EVPV*LPVSPPLPRPLQKMI KDKKEVGFFQSIQALMQTC\GEKVMADI QDLFRFLQLLCEGHNNDFQNYLRTQTGI INIIICTVDYLLRLQESI 39	RGKG
KDKKEVGFFQSIQALMQTC\GEKVMADDQDLFRFLQLLCEGHNNDFQNYLRTQTGINIIICTVDYLLRLQESI	
QDLFRFLQLLCEGHNNDFQNYLRTQTGI INIIICTVDYLLRLQESI 39 1389 A 767 I 1030 TLDLTGPLLLGGVPNVPKDFRGRNRQFC RNLSVDGKNVDMAGFIANNGTREGCAA FCDGRRRQNGGTCVNRWNMYLCECPLI KNCEQGEWPASSIPPVTAAWEALLLDVI VRGLHIQVRQPLVVYAAFTVDSHRPLQI RRAPAPASGVPSPSGVGWDR*AGPAEPS ATVIISVPWYLGLMFTRKEDSVLMEA PTSFRLQVTGAPCHQGTC*VGARGRDPN LRVTDGEWHHLLIELKNVKEDSEMKHL TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGEA QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*QVILKFTW*E*GAKITILRE RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	
INIIICTVDYLLRLQESI 39 1389 A 767 I 1030 TLDLTGPLLLGGVPNVPKDFRGRNRQFG RNLSVDGKNVDMAGFIANNGTREGCAA FCDGRRRQNGGTCVNRWNMYLCECPLI KNCEQGEWPASSIPPVTAAWEALLLDVI VRGLHIQVRQPLVVYAAFTVDSHRPLQI RRAPAPASGVPSPSGVGWDR*AGPAEPS ATVIISVPWYLGLMFRTRKEDSVLMEA PTSFRLQVTGAPCHQGTC*VGARGRDPN LRVTDGEWHHLLIELKNVKEDSEMKHL TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGEA QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*QIVILKFTW*E*GAKITILRE RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	
39 1389 A 767 I 1030 TLDLTGPLLLGGVPNVPKDFRGRNRQFG RNLSVDGKNVDMAGFIANNGTREGCAA FCDGRRRQNGGTCVNRWNMYLCECPLI KNCEQGEWPASSIPPVTAAWEALLLDVI VRGLHIQVRQPLVVYAAFTVDSHRPLQI RRAPAPASGVPSPSGVGWDR*AGPAEPS ATVIISVPWYLGLMFRTRKEDSVLMEA PTSFRLQVTGAPCHQGTC*VGARGRDPN LRVTDGEWHHLLIELKNVKEDSEMKHN TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGEA QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*QIVILKFTW*E*GAKITILRE RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	
RNLSVDGKNVDMAGFIANNGTREGCAA FCDGRRRQNGGTCVNRWNMYLCECPLI KNCEQGEWPASSIPPVTAAWEALLLDVI VRGLHIQVRQPLVVYAAFTVDSHRPLQI RRAPAPASGVPSPSGVGWDR*AGPAEPS ATVIISVPWYLGLMFRTRKEDSVLMEA' PTSFRLQVTGAPCHQGTC*VGARGRDPN LRVTDGEWHHLLIEKNVKEDSEMKHN TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGEA QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*QIVILKFTW*E*GAKITILRE RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	GGCM
KNCEQGEWPASSIPPVTAAWEALLLDVI VRGLHIQVRQPLVVYAAFTVDSHRPLQI RRAPAPASGVPSPSGVGWDR*AGPAEPS ATVIISVPWYLGLMFRTRIKEDSVLMEA' PTSFRLQVTGAPCHQGTC*VGARGRDPN LRVTDGEWHHLLIELKNVKEDSEMKHL TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGEA QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*QIVILKFTW*E*GAKITILRE RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	
VRGLHIQVRQPLVVYAAFTVDSHRPLQI RRAPAPASGVPSPSGVGWDR*AGPAEPS ATVIISVPWYLGLMFRTRKEDSVLMEA' PTSFRLQVTGAPCHQGTC*VGARGRDPN LRVTDGEWHHLLIELKNVKEDSEMKHL TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGEA QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*QVILKFTW*E*GAKITILRE RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	RFGG
RRAPAPASGVPSPSGVGWDR*AGPAEPS ATVIISVPWYLGLMFRTRKEDSVLMEA' PTSFRLQVTGAPCHQGTC*VGARGRDPM LRVTDGEWHHLLIELKNVKEDSEMKHL TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGEA QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*QVILKFTW*E*GAKITILRE RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	PGTT
ATVIISVPWYLGLMFRTR\KEDSVLMEA\ PTSFRLQVTGAPCHQGTC*VGARGRDPN LRVTDGEWHHLLIELKNVKEDSEMKHL TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGE\ QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*Q\VILKFTW*E*GAKITILR\ RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	ETVL
PTSFRLQVTGAPCHQGTC*VGARGRDPN LRVTDGEWHHLLIELKNVKEDSEMKHL TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGE, QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*Q\VILKFTW*E*GAKITILRE RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	SPSTP
LRVTDGEWHHLLIELKNVKEDSEMKHL TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGE/QAAPRL 40	
TLDYGMDQVSWHLHLLWG*TLPPAQGI SEDKVSVRRGFRGCMQVRGGCGGRGE/QAAPRL 40 1390 A 801 69 399 IHKIIIHKEDLNKWKYILCSGMERLSTVN PQIIYKFNA*Q\VILKFTW*E*GAKITILRE RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	
SEDKVSVRRGFRGCMQVRGGCGGRGEA QAAPRL	
QAAPRL	
40 1390 A 801 69 399	ACPS
PQIIYKFNA*Q\VILKFTW*E*GAKITILRE RGLVLVPLSTC*VKYLLDKVLPHIKTYY VNKSVVLVQVTIM	(TTO V 174)
RGLVLVPLSTC*VKYLLDKVLPHIKTYY	
VNKSVVLVQVTIM 41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	
41 1391 A 835 7 195 SMLKERKVFQFPSCLFFQYITWLGPPYH	ĖAK
	NA EL
<u></u>	
42 1392 A 841 1 415 GSTHASGYDKTPDFILQVPVAVEGHIIHV KASFGDECSHHAYLHDQFWSYWNSLKI	
QGIGTVASNLSQL*TLNAPFPELLLFRSL	
FVLT*\RFGPGLVIYWYGFIQELDCNRER	
KACFPTNIVTL	CILL
43 1393 A 845 358 92 PALSPAPVPQKKGSPLPLDPCLGPSSWLI	ISVG
43 1393 A 843 338 92 FALSPAP V PORROSPEPED PED PED PED PED PED PED PED PED PE	
QRPMLPPSHAGLARPPPPEPISVP	
44 1394 A 853 452 1 LPQYCFFPRLSPKSKLVKHSAL**PSALK	TTQQ
44 1394 A 853 452 1 EPQTCFFFRESFRSRLVRISAL TSALR	
PPTHMLRSASQPLNQAPTLVKGHPPSRF	
QVSCPPQPTLPREKPLPLHLRPPPRPAQP	
PLTFSTRRNVDPEIPERFR	
45 1395 A 894 379 162 GVYPPTVFDNYSVQTSVDGQIVSLNTW	
QEEYD/RLRTLS*PQTSIFVICFSIGNLEFF	'DTAG
WLSMSMGK	DTAG PIYGT
46 1396 A 900 1 366 TTKKTLISNNVSSRSLPILPELKAFSLAFY	DTAG PIYGT
EIQKYMRT/DQ*CVTHDISLYIVTKLALI	PIYGT
VFLFHQLNIT**CLHFFTMTTFIAIPFSFLI	PIYGT NDPL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion D/KSLAMLPRLVSNSWPQVILPP QLQNLASRGCL*SQLLRRLRRENRLNPGGGG CSEIAP\CTPAWVTQRDFFRKKK
48	1398	A	963	216	308	HFTPDRIAIVKNTRDSHCWRGC*EEGAPARC
49	1399	A	967	466	1	PRKRESWWGERLP/PRGFPPAAEDAPAPGWK GRKHASRTARAHVFHPIRQSIRSPVRGRPGDP RAAHTRSAGTRLQCKASRGG*GKGPAPTR*E GGPGSAPAPLPASSGCSLFPDSSPWTPPPPAPG AAAAQP**TPRCPAALRAGAHIGRVGRPY
50	1400	A	973	45	421	EKCIQALDVFVFCYIDHSSHCLMSCD*E/DQA LNFMPLEMEPKMSKLAFGCQRSSTSDDDSGC ALEEYAWVPPGLRPEQIQLYFACLPEEKVPY VNSPGEKHRIKQLLYQLPPHDNEVRYCQSLSE E
51	1401	A	992	2095	194	IRIRHEAARSCLGCAAGHVPAPGLRLLPTVRG PPGRRGPAAPGCVCY*SGESTFVSHVPQRMA WPGSAPPRGFHPLQSQTSPSDTVSSPQLSKEE DGPGWEHPLSSSL*SLGQAGGNH*QPEELAG WEPRGPPSLAPSSPT/TMWTALVLIWIFSLSLS ESHAASNDPRNFVPNKMWKGLVKRNASVET VDNKTSEDVTMAAASPVTLTKGTSAAHLNS MEVTTEDTSRTDVSEPATSGVAADGVTSIAPT AVASSTTAASITTAASSMTVASSAPTTAASST TVASIAPTTAASSMTAASSTPMTLALPAPTST STGRTPSTTATGHPSLSTALAQVPKSSALPRT ATLATLATRAQTVATTANTSSPMSTRPSPSKH MPSDTAASPVPPMRPQAQGPISQVSVDQPVV NTTNKSTPMPSNTTPEPAPTPTVVTTTKAQAR EPTASPVPVPHTSPIPEMEAMSPTTQPSPMPYT QRAAGPGTSQAPEQVETEATPGTDSTGPTPRS SGGTKMPATDSCQPSTQGYMV/DHH*APHP GRGRQNSPSGGAVTRGDPFHHSLGFVCPAGL *ELQEEGLHPGGLLNQRDVCGLRNVRGAGA WREAWPLPRPFLLPLRPNQVLPNSFGAIEEIC QMLKHI
52	1402	A	994	1	462	ESGEFLVSFTLKKPTNVFHHINGMKFFNK/LIF *SHTDIAFYKIQHPFMLKALTKWA*EGT*PDR RYLH*SLRLNGEQLKTFPLRSGMR*G/CAILPL VLNAMLSIVPAVVPAGKTRHEKEITCPLIGQE EK*FS*FVGDMNTCVENKKESKKLLE
53	1403	A	1011	1	630	PEVIQQSAYDSKADIWSLGITAIELAKGEPPNS DMHPMRVLFLIPKNNPPTHCWRRŁLESFKEV *LMLA*TKDPSI\RPTAKELLKHKFIVKNSKKT SYLTELIDRFKRWKAEGHSDDESDSEGSDSES TSRENNTHPEWSFTTVRKKPDPKKVQNGAEQ DLVQTLSCLSMIITPAFAELKQQDENNASRNQ AIEELEKSIAVAEAAGPG
54	1404	A	1016	1	222	ISIDA*KAFDKIQH/CFMITTLKKLGIDGKYLN TIKAIDDRHTVSTILNVEKLKAFL*RSGTRQRF PISGSGARI
55	1405	A	1033	3	366	HASVDGDEGSDDVYYYYTPAILRELQALNTA EAAEHRPEEDRMLSEDPWRPAHMIKGYMPL HNIPHTEVIDVTGLNQSHLYQHLNKGTPMKT QKRAA\LYTWHVLEQLEILRQINQQSHGPG
56	1406	A	1044	5	429	SVLTLQTRSPSKPLS\RKLMDWEVVSRNSISE DRLETQSRASRSPPVTPNQSQETPVDGKPLAL PPNQSQKNIRYHIHYLHLQYYLDRHISATLPIP SSSGIPTPIAVITDALTDLVELILGQPCSEESGR APGTLFLLAL

		·	T-500	F 17 . 1		LA
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide]	in	location		I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	1	USSN		corresponding	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi ng to first	to last amino acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence		1	914	amino acid		T=Threonine, V=Valine, W=Tryptophan,
		Į	1	residue of	of peptide sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	Ĭ	1	{	peptide	Seductice	/=possible nucleotide deletion, \=possible
	ł	l	1	sequence		nucleotide insertion
57	1407	A	1050	11	430	GAYAFETNGFPIMLVLTTDKIEGDVGIAGLYD
)))	1407	A	1,050	**	450	MH\ISLPMAFLLRTLVRCTSYIIPVTHVLSTPV
·	i	1	1	[TCLRRREKDGVIVDVLSDTASNHNGFPVEEH
	Į		Ì]		ADDTHPARLOGPTLRSQPMGPLKHKAFEERA
		1			ĺ	NLGLVORRLRLED
58	1408	A	1058	258	419	LKHRDTPVVGANNRALSCTPLTSLTLCALCPL
50	1400	1	1030	220	''	PCLGCPTXATCRLYQTTVAVVF
59	1409	A	1064	3	425	KAFSFTTSLIGHQRMHTGERPYKCKECGKTF
)	140)	1.	100.	*		KGSSSLNNHQRIHTGEKPYKCNECGRAFSQC
	i		1	i	•	SSLIOHHRIHTGEKPYECTQCGKAFTSISRLSR
	1	Ì	1	1	ł	HHRIHTGEKPFHCNECGKVFSYHSALIIHQRIH
	!		ļ	l		TGEKPYACKDVGK
60	1410	A	1065	204	419	GGPPGPFLAHTHAGLQAPGPLLAPAGDEGDL
••				ļ	{	LLLAVQQSCLADHLLTASWGGK/DPIPTKALG
		1	1		1	EGQEGLPLTV
61	1411	A	1079	3	383	RHSRAHLCQPFHLVMRDLLQLGQDIPQGCHY
						LEENHLIHRDIAARNCLLSCAAPTRAATIGDF
						GMARYIYRTRYYQLGDRAL/LPRKWMPPEAL
	1	1		Ì	İ	LEGIFTYNTDSWTFGVLLWEIFSLGYMPYPGR
		1		_		TN
62	1412	A	1080	1	859	VVEFLWSRRPSGSSDPRPRRPASKCQMMEER
į.	1					ANLMHMMKLSIKVLLQSALSLGRSLDADHA
ŀ		1.			į	PLQQFFVVMEHCLKHGLKVKKSFIGQNKSFF
ĺ	(1	İ	1	1	GPLELVEKLCPEASDIATSVRNLPELKTAVGR
	ŀ	1	1			GRAWLYLALMQKKLADYLKVLIDNKHLLSE
ĺ	ļ					FYEPEALMMEEEGMVIVGLLVGLNVLDANL\
1	1	ł	1	ł	Ì	CLKGEDLDSQVGVIDFSLYLKDVQDLDGGKE
Ì	ł		1		ļ	HERITDVLDQKNYVEELNRHLSCTVGDLQTK
1			[IDGLEKTNSKLQERVSAATDRICSLQEEQQQL
	1	 	1.000			REQUELIR
63	1413	A	1083	2	615	SSFAKHKRIHTGEKPFICLECGKAFTSSTTLTK
	1			ĺ	Ì	HRRIHTGEKPYTCEECGKAFRQSAILYVHRRI HTGEKPYTCGECGKTFRQSANLYAHKKIHTG
	}					EKPYTCGDCGKTFRQSANLYAHKKIHTG\EKP
}					e	YKCKECGKAFKSYYSILKHKRTHTRGMSYEG
1	1	1		Ì	ľ	DEC/QRSLN/RSSILSNHKIIHNEEK/PLKCEKCE
		1	1			KAFNHTSICCRHKKN
	1414	1_	1004	946	1	KKQDLSSSLTDDSKNAQAPLALTESHLATLA
64	1414	A	1084	740	1 1	SSSQSPEAIKQLLDSGLPSLLVRSLASFCFSHIS
	1	1	1	1		SSESIAQSIDISQDKLRRHHVPQQCNKMPITAD
į	ł	1		[LVAPILRFLTEVGNSHIMKDWLGGSEVNPLW
1	-	1	1]		TALLFLLCHSGSTSGS\HNLG\AQQDQCKISFS
[1	1		!		FFSWLTTGLTTQQRTAIE\NATVAFF\LQC\\SC
1	ł	1		1		HPNNQKLMAQVLCELFQTSPQRGNLPTSGNI
1	1	1	1 .	1		S\GFIR\RLFLQLMLEDEKVTMFLQSPCPLYKG
ĺ	1	1		[1	RINATSHVIQHP\MYGAGHKFRTLHLPVSTTL
	[1	1		j	SDVLDRVSDTPSITAKLISKQKDDKKKK
65	1415	A	1087	103	324	PRAFEFVHTEMIVG/RVQNIHLFTLQVLEDRA
1 00	1417	1^	1007	1.55	1	LFTMSVGSSLWSTYLIHVMALP/DRELLKPNA
}	1	1		}		SVALHKLSNALV
66	1416	A	1095	3	493	HETCSVTHIVSFSLPFLNPSHPASTPGHTENEQ
00	****	1"	1000	١	""	PSLVWFDRGKFYLTFEGSSRGPSPLTMGAQD
1	1	İ		1		TLPVAAAFTETVNAYFKGADPSKCIVKITGE
1		1	İ	1		MVLSFPAGITRHFANNPSPAALTFRVINFSRLE
1		i				HVLPNPQLLCCDNTQNDANTK\EFWVNMPNL
		1		1 .		MTHLK
67	1417	A	1098	57	356	LKLTSLGFIIGVSVVGNLLISILLVKDKTLHRA
"	1 ****	1	1.570	1	1	PYYFLLDLCCSDILRSAICFPFVFNSVKNGST
i		-	1	1	1	WTYGTLTCKVIAFLGVLSCFHTAFMLFCISVT
						

SEQ ID NO; of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	Į	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		Ì	ĺ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
İ	}	}	ļ	sequence	1	nucleotide insertion
		 	 	soquenes	 	RYL
68	1418	A	1106	1	1326	MGKISATGINMGTKCSWALVWHLESYDPKH
		1	****	· ·	.520	YEREGMQDWKTASGQSEEATQQSSQKPQPH
	İ	}	1	İ	1	YTTYQSSSFLKYSSESHLLAWRENSSEGSFQF
	ł	ł	{	}	ł	PGRSRARPPRTRQQRRGAAAGPGRGAVRLG
		1				HPQSAAQPQLRAAARIPESPAAFPAQPRPGSA
			1	•		RNSDASGPASLSRTLGRASSPRPPQAPDVTAP
	İ	1				SPAALAPRAARGGSRAAALAGAEAEEPLRTL
	[ĺ		(APRPTRAAAPPPPPPPPPPLPPGAPPPPVRCVSR
		Ì				RARAPPWR/PAATGPPP\RPVAPSRKLGSARAP
		1	1			APALQIRKGTSSGLPGRGGGSGPGNNLSSVA
	}	}	1		į	GNWRGSSFAVERPGMAKYQGEVQSLKLDDD
		1				SVIEGVSDQVLVAVVVSFALIATLVYALFRNV
	ŀ	[HQNIHPENQELVRVLREQLQTEQDAPAATRQ
		İ				QFYTDMYCPICLHQASFPVETNCGHLFCGSLT
		1	1			PNSIW
69	1419	A	1107	2	466	
0,5	1417	^	1107	-	400	FDTARLHEFGTSITQIFAVDNREDLQKWMEA
	Ì					FWQHFFDLSQWKHCCEELMKIEIMSPRKPPLF
	(ĺ			LTKEATSVYHDMSIDSPMKLESLTDIIQKKIEE
	1					TNGQFLIGQREESLP/SS/CGPHSLMVTIKWSS
70	1420	A -	1111	698	23	RKRY/SYPASEPLHDEKGKKRQAPLPPSDK
70	1420	Α .	1111	098	23	ALRRLHYVRATKVFLSFRRPFWREEHIEGGH
						SNTDRPSRMIFYPPPREGALLLASYTWSDAAA
			{			AFAGLSREEALRLALDDVAALHGPVVRQLW
	ĺ	İ				DGTGVVKRWAEDQHSQGGFVVQPPALWQT
			1			EKDDWTVPYGRIYFAGEHTAYPHGWVETAV
		1	1			KSALRAAIKINSRKGPASDTASPEGHASDMEG
	}		1			QGHVHGVASSPSHDLAKEEGSHPPVQGQLSL
71	1421	A	1119	2	385	QNTTHTRTSH
'	1421	^	1117	2	202	QKQTLQNGYLDSSMDILYLGSLPPELQVSSDE
	ł		}			PPGPPEQAGLSQFHLEPETQNPETTEEIQSS\LQ
	}	ł	1			QEAAAQLPQLPEVVELSSTKA\EAPALPSQSL EGVHSSTEQKAPAQQLPAFEEILAPLLIHHE
72	1422	A	1127	1	906	
12	1422		112/	1	906	HAQYVGPYRLEKTLGKGQTGLVKLGVHCIT
	l .	ļ				GQKVAIKIVNREKLSESVLMKVEREIAIL\RLI
	·					EHPHVLKLHGVYENKKYFPPDELTSGPSMLA
]	ļ			QVSPHGKLSARRSWDLLSGFPRYLVLEHVSG
	}	1				GELFDYLVKKGRLTPKEARKFFRQIVSALDFC
	}	ļ	!			HSYSICHROLKPENLLLDEKNNIRIADFGMAS
		i]			LQVGDSLLETSCGSPHYACPEVIKGEKYDGR
		1	[RADMWSCGVILFALLVGALPFDDDNLRQLLE
	1]			KVKRGVFHMPHFIPPDCQSLLRGMIEVEPEKR
73	1423	A	1128			LSLEQIQKHPWYLGGNFIS
15	1443	^	1128	1	802	LRNALDVLHREVPRVLVNLVDFLNPTIMRQV
		1				FLGNPDKCPVQQA/MLEPLGSKTETLDLRAE
		({			MPITCPTQNEPFLRTPRNSNYTYPIKPAIENWG
		1				SDFLCTEWKASNSVPTSVHQLRPADIKVVAA
	}	ł	!			LGDSLTTAVGARPNNSSDLPTSWRGLSWSIG
]			GDGNLETHTTLPNILKKFNPYLLGFSTSTWEG
		l '	[]			TAGLNVAAEGARARDMPAQAWDLVERMKN
						SPDINLEKDWKLVTLFIGGNDLCHYCENPEA
74	1404	<u> </u>	1122		400	HLATEYVQHIQQALDILSE
74	1424	A	1139	60	480	FREPCLLVPGDHQPLREASWLA/LPPIGLWGT
						DSPLCCVEVAIPCNKGAHSVGLKGWLLAQG
		·				VLGMRDTIPQEHPWESTPDLCFCRDPEEIEVE
			[EQPAADAAVAKGEF/QGEQIAPVPA\IIAAHPE
		L				AADPAPVHTTAHPKGA
75	1425	A	1147	2	413	PFPHQHPQEP'KGSCWPQSALRGQCPGPVLGV TTTSDLCSLQVPVSSHRNPLLDLAAYDQEGR

						
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	Ì	1	ł	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł	ł	l	ŀ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ	ì	1	peptide		/=possible nucleotide deletion, \=possible
		1	ļ	sequence		nucleotide insertion
	ļ					RFDNFSSLSIQWESTRPVLASIEPELPMQLVSQ
				ł	ļ	DDESGQKKLHGLQAILVHEASGTTAITATAT
ŀ	[ł	ł	ł .	GYQESHLSSAR
76	1426	A	1155	38	410	PIISAPAQDDPILLSFIHCLHANLLCVWRRDVK
, ,	1 120		1	**		PDCKEIWIFWWGDEPNLV\VOYIMNCMLWK
	f	1	J	j	ļ	KDSGKMAFPMNVGRC/FFKEIHNLLERCLMD
ļ			1	ĺ		KNFVLIGKWFVRPYYKDEKPVNKSEHLSCAF
i	1	l	l .			T
77	1427	A	1162	526	350	RFPQGLEDVSTYPVLIEELLSRGWSEEELQGV
\	1427	A	1102	320	330	
70	1400	 -,	1,,,,,		1202	LRGNLLRVFRQVEKVQEENKWQSPLED
78	1428	Ā	1171	1	1293	MAESASPPSSSAAAPAAEPGVTTEQPGPRSPP
				[İ	SSPPGLEEPLDGADPHVPHPDLAPIAFFCLRQT
		1	1			TSPRNWCIKMVCNPWFECVSMLVILLNCVTL
	Į.	1	ł	ł		GMYQPCDDMDCLSDRCKILQVFDDFIFIFFA
1			1	1		MEMVLKMVALGIFGKKCYLGDTWNRLDFFI
				J	}	VMAGMVEYSLDLQNINLSAIRTVRVLRPLKA
		1	1	Į		INRVPSMRILVNLLLDTLPMLGNVLLLCFFVF
1						FIFGIIGVQLWAGLLRNRCFLEENFTIQGDVAL
ł	ŧ	l	ł	ł	ł	PP\YYQPEEDDEMPFICSLSGDNGIMGCHEIPP
}	į	l	ł	1	}	LKEQGRECCLSKDDVYDFGAERQDLNASGL
	1	1	1		1	CVNWNRYYNVCRTGSANPHKGAINFDNIGY
l						AWIVIFQVITLEGWVEIMYYVMDAHSFYNFI
í	İ	1	ĺ	Í	ĺ	YFILLIIVSVREPGLLGGSFSTAQSPKCQGDSFP
					}	GVAAESLLLRGWVLWLPGGG
79	1429	A	1175	1	405	PNDFFKDMFPDLPGGPLGPIKAENDYGAYLN
1 .	1.2	1	***	ļ [*]	102	FLSATHLGGLFPPWPLVEERKLKPKASQQCPI
	ŀ	1	ì	{	į.	CHKVIMGAGKLPRHMRTHTGEKPYMCTICE
	ļ	İ	I			VRFTRQDKLKIHMRKHTGERPYLCIHCNAKF
1		{	1			VHNYDLKNHMR
80	1430	A	1182	25	198	EMNELSQQLSQQGGRGASQCPSPPAPTLPNPT
00	1430	^	1102	23	170	PLCQLQRVNTGLPTPPCHPGAGAA
01	1431	 	1106	254	583	KTVLDVGAGTGILSIFCAQAGARRVYAVEAS
81	1431	A	1186	254	283	
1	(1	1			AIWQQAREVVRFNGLEDRVHVLPGPVETVEL
1	[ł		ł		PEQVDAIVSEWMGYGLLHESMLSSVLHARTK
		ļ	1			VVKDGGFFLPXSSELFM
82	1432	A	1187	2	716	DFVDAARNLPLESTKSPAEPSKSVPSLE\DPRA
1 -	ľ					SSQGLPSQGPVQNQGRRGEQRPKKF/TVIQHT
	!					SSFEKSDSLEQPSGLEGEDKPLAQFPSPPPAPH
	1		[{	{	GRSAHSLQPKLVRQPNIQVPEILVTEEPDRPD
		1	1	1		TEPEPPPKEPEKTEEFQWPQGSQTLAQFPVEK
1	ļ	1	1	1	1	LPPKKKRLGLAKMAQSSGESSFESSVPLFRSP
1]	1	}	1		SQESNVSLSGSSRSALFERDDHGKAEAPSPSF
1			1			DMGPKPLGTHMLTV
83	1433	A	1188	517	804	ESPGLSKVLRTGAFAYPFLFDNLPLFYRLGLC
	1			[-	WGRGHGCGQEALSTSHGYHLFCALLTGFLFA
	1			1	1	SHLPERLAPGRFDYIGHSHQLFHICAVLGTHF
	}	1		1	1	0
84	1434	A	1192	45	476	LGDVGFWVERTPVHEAAORGESLOLOQLIES
ا ا		Α	11172	**	1,0	GACVNQVTVDSITPLHAASLQGQARCVQLLL
İ		1	1	I	1	AAGAQVDARNIDGSTPLCECLRLGQHRVCEA
ĺ		1	1	1		LAVLRGQGQPSPVHSVPPARGLHXREFRMC*
	1	1		1		GFLFDVGXNLEAHEFHFGEP
05	1425	 	1104	160	420	
85	1435	A	1194	69	410	KRSEEASAPPFPLGGTGAAPTRASLPEQILLPR
}	}	İ			}	SCLEARKSQPDEKLLSALHNSRTWN*EPRRSQ
[!	[HRLVSPEVHPGRRGSSPGVAECKLTSAYFRT
	<u> </u>	<u> </u>	L.	<u> </u>	L	GRSPCPSLPGTTRTNSLL
86	1436	A	1215	3	405	LPSHTCGNPGRLPNGIQQGSTFNLGDKVRYSC
l		1		l	I	NLGFFLEGHAVLTCHAGSENSATWDFPLPSC
	1	1	I	1	I	RADDACGGTLRG/AEWHHLQPPLPLG/ATKN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion NADCTWTILAELGDTIALVFIDFQLEDGYDFL EVTGTEGSSLW GTARFGPMVGFGANRRAGRLPSLVLGVLLV
		·				VIVVLAFNYWSISSRHVLLQEEVAELQGQVQ RTEVARGRLEKRNSDLFAVVGHAQETDRPEG GRLRPPQQPAAGQRGPREEM\EDDKVKLQNN ISYQMADIHHLKEQLAELRQEFLRQEDQLQD YRKNNTYLVKRLEYESFQCGQQMKELRAQH EENIKKLADQFLEEQKQETQKIQSNDGKELDI NNQVVPKNIPKVAENVADKNEEPSSNHIPHG
88	1438	A	1218	1	534	PEFGTTISCGYLMATDVSRRPSVHKAVEIEQE RVKSAGAWIIHPYSDFRFYWDLIMLLLMVGN LIVLPVGITFFKEENSPPWIVFNVLSDTFFLLD LVLNFRTGIVVEEGAEILLAPRAIRTRYLRTW FLVDLISSIPVDYIFLVVELEPRLDAEVYKTAR ALRIVRFTKILSLLRL
89	1439	A	1223	1	743	MGFDEVFMINLRRRQDRRERMLRALQAQEIE CRLVEAVDGKVGMLTRSNAAPGRHLAMLET LVVVAPRFVDADNLILNPDTLSLLIAENKTVV APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNL VAFYPPHPDYTWSFDDIIVFAFSCKQVAEVQMY VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ LEVMVPSSPSSAQSMAVVSADHIGLVISYL
90	1440	A	1227	2	349	NKTSFIFYLKNIVVADLIMTLTFPFRIVHDAGF GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/ YDRY/WKVVRHL/WDSWMTGI/SFTRVYLLG LGARLVWFGKLILAKGGHGGISWL
91	1441	A	1245	3		LGSSDVRAPQRSELGAESPSRMVASQAYNLT SALTPILTRSRVLNEEPLTLAGF\SRAPANLSD VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF QTQAGAQIPIERLASERAITVKVPNNSDWAAR GHRSSANSVVQPQAFVGAVVTLDSSNPAAV LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR PNEHNCSASRRIRPESLQGADHRPYTFFISPGT RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD LQTARSTFFLVNDWLSVETEANGGLVEKEVL AASKASFRVPTPS\AALLRFRRLLVAELQRGF FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL FLGANAVWYGAVGDSAYSTGRVSRLNPLSV DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV GWGWGPGSTGNGAWASAPCPEPPLSSAAAR GKGVHQRLLGKGQHT
92	1442	A	1246	5	562	VFDEENILNELNDPLREEIVNFNCRKLVATMP LFANADPNFVTAMLSKLRFEVFQPGDYIIREG AVGKKMYFIQHGVAGVITKSSKEMKLTDGS YFGEICLLTKGRRTASVRADTYCRLYSLSVD NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN SILLQKFQKDLNTGVFNNQENEILKQIVKH
93	1443	A	1249	180	901	TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT ERPGAERPSLLPNGKENSSGTPRVPPASPSSHS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LAPPSGERSRLARGSTIRSTFHGGQVRDRRAG GGGGGGVQNGPPASPTLAHEAAPLPAGRPRP
94	1444	A	1261	3	385	TTNLFTKLTSKLTRRVADEPERIGGPEVTRRP RQEDHLSPGGRGCSEL KFSQWGLTKPKLSNASP/WISLVKKLMKKWS
94	1444	A	1201	3	363	VTQNLTFREQLEAGIRYFDLRVSSKPGDADQ EIYFIHGLFGIKVWDGLMEIDSFLTQHPQEIIFL DFNHFYAMDETHHKCLVLRIQEAFGNKLCPA CR
95	1445	A	1282	2	550	GPRDNPGVEDPRFEIVEHFGIAWFTFELVARFA VAPDFLKFFKNALNLIDLMSIVPFYITLVVNL VVESTPTLANLGRVAQVLRLMRIFRILKLARH STGLRSLGATLKYSYKEVGLLLLYLSVGISIFS VVAYTIEKEENEGLATIPACWWWATVSMTT VGYGDVVPGTTAGKLTASACILA
96	1446	A	1294		1456	QLLPPSNRENAGLLVGRCLCSAALRPVGDLIT SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIQ YFKKFFDANCNEKDYNPVAAGQGQETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMIF/LV/LVILLGSLYLVTLILAV/VAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRRKKRKQKEQSGEEKDED EFQKSESEDSIRRKGFFFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE
97	1447	A	1295	2	2057	IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HILKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR
98	1448	A	1304	118	453	SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS
99	1449	A	1306	3	1660	CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE TGTGTAYEGFLSVPRPSGVRRGWQRVFAALS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DSRLLLFDAPDLRLSPPSGALLQVLDLRDPQF SATPVLASDVIHAQSRDLPRIFRVTTSQLAVPP TTCTVLLLAESEGERERWLQVLGELQRLLLD ARPRPRPVYTLKEAYDNGLPLLPHTLCAAILD QDRLALGTEEGLFVIHLRSNDIFQVGECRRVQ QLTLSPSAGLLVVLCGRGPSVRLFALAELENI EVEVPKIPESRGCQVLAAGSILQARTPVLCVA VKRQVLCYQLGPGPGPWQRRIRELQAPATVQ
					·	SLGLLGDRLCVGAAGGFALYPLLNEAAPLAL GAGLVPEELPPSRGGLGEALGAVELSLSEFLL LFTTAGIYVDGAGRKSRGHELLWPAAPMGW GYAAPYLTVFSENSIDVFDVRRAEWVQTVPL KK\VRPLNPEGSLFLYGTEKVRLTYLRNQLAE KDEFDIPDLTDNSRRQLFRTKSKRRFFFRVSE EQQKQQRREMLKDPFVRSKLISPPTNFNHLV HVGPANGRPGARDKSP
100	1450	A	1318	918	190	SLCVPGPVDTGTFAVMSVMVGSVTESLAPQA LNDSMINETARDAARVQVASTLSVLVGLFQV GLGLIHFGFVVTYLSEPLVRGYTTAAAVQVF VSQLKYVFGLHLSSHSGPLSLIYTVLEVCWKL PQSKVGTVVTAAVAGVVLVVVKLLNDKLQQ QLPMPIPGELLTLIGATGISYGMGLKHRFEAG\ PPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGK IFALRHGYRVDSNQVWVMRDV
101	1451	A	1353	220	445	DWPDLFTYPLIGSPKCFQSARPE\RMYRRTVR SSHGNHALQEVLPRSGHGTEFTKQKHLEAAD HGHPPARMSIFSR
102	1452	A	1363	542	2	AHLLMLNLAL\TDLL\YLTSLPFLIHYYASGEN WIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRY CVIIHPMSCFSIHKTRCAVVACAVVWIISLVA VIPMTFLITSTNRTNRSACLDLTSSDELNTIKW YNLILTA\LLCLPLVIVTLCYTTIIHTLTHGHAN \DSCLKQKARRLTILLL
103	1453	A	1371	2	410	CHSTESSSDFILPGDYLLGGLCPLHSGCLQVC SFNEHGYHLFQAMRLAVEEINNSTALLPNITL GYQLYDVCSDSANVYATLRVLSLPGQHHIEL QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP FLVPMLLEQ
104	1454	A	1376	3	432	NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF
105	1455	A	1379	2	396	GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA
106	1456	A	1383	1	432	EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP
107	1457	A	1386	719	558	FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL
108	1458	A	1397	631	2	RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK AYVCSLLASQLLIIPVFAIGGANVWVLGLLLF

NO. of No. of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Docation	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glucine Acid,
	1	1 4 4					
146	,						
minio acid residue of peptide residue of peptide sequence residue of peptide sequence residue of peptide sequence residue of peptide sequence residue of peptide resi		uence			correspondi	1.5	
residue of peptide pep	uence			914			
		-					
	ĺ	ĺ	[ĺ		sequence	
109	İ					ĺ	
109		 			sequence		
LSPSLITEVVILRNRRPGKSLVR		ļ				i	
109)	1			ļ]	-
WSHNSNSMCWGRQCPYSGCKEALIRTDOM RYTSRKAYKTIGGTPRGPWSTITINPSESDS GVYCCRIEVPGWPNDVKINNTINLORASTI HEDISSLITRGSGNGEERGICKKUSIGLDWM LAELAFPVGVLATCA*SLLSC*YCVILFPCSCF FFHSPDALFSILLISCYPFSYCFFYTLFFSSCF CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLLASSPPFLIFLASI CLLASSPPFLIFLASI	100	1450	 	1402	18	207	
RYTSRKSAKYRLQGTIPRGDVSLTILINPSSIDS	109	1439	^	1402	13	307	
110		1	1		-	!	1 1 1 1
110	1		1			ĺ	
LAELAFPYGVLATCA*SLLSC*YGVLFPCSCF FFHSPDALFSLLLSCYFPSYCFFYYLFFSSSP.	110	1460	_	1/21	3	350	
	110	1400	Α	1421	13	330	
CLLLASSPPLFILLASL	ļ						i i
111			}	Ì	Į.	}	
QCALKPDLSYLNNSSSSSTPATSAGGGIGGS TSSSNPPVATVFQQSDSYGFVNTAESST TSSSNPPVATVFQQSDSYGFVNTAESST SDSLLFSQDSKLATTS TTSWTTSCTT*SGASSGFGWTPTTWWR SRRSSQRTCSRACSGAWSRTW*RSS*TSSSSC STSCSSSSSRSCGRPGGFLGARGYHITSCLNSC MSSTTSSTTSTF HEDMTHYDRLVDE*ALNAGKQRYEKMISG MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNILIDFTKKGFLLRGQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNITHQQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNITHQQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNITHQQISEMLKTR GIFLTFLLSNFLUCVLLFYVSFYLFQSCNFVL MYLGEIVRNITHQQISEMLKTR MYLGEIVRNITHQQISEMLKTR MYLGEIVRNITHQQISEMLKTR MYLGEIVRNITHQQISEMLKTR MYLGEIVRNITHQQISEMLKTR MYLGEIVRNITHQQISEMLKTR MYLGEIVRNITHQQISEMLKTR MYLGEIVRNITHQQISEMLKTR MYLGEIVRNITHQQISEMLKTR MYLGEIVRNITHQASSPTNNIFASSK MYLGEISSPTYNDFVCMPLUGNKKRYLTHQQI MYLGEIVRNITHQASSPTNNIFASSK MYLGEISSPTYNDFVCMPLUGNKKRYLTHQQI MYLGEIVRNITHQASSPTNNIFASSK MYLGEISSPTYNDFVCMPLUGNKKRYLTHQQI MYLGEIVRNITHQASSPTNNIFASSK MYLGEISSPTYNDFVCMPLUGNKKRYLTHQQI MYLGEIVRNITHQASSPTNNIFASSK MYLGEIPROV MYLGEIVRNITHQASSPTNNIFASSK MYLGEIPROV MYLGEIVRNITHQASSPTNNIFASSK MYLGEIPROV MYLGEIVRNITHQASSPTNNIFASSK MYLGEIPROV MYLGEIVRNITHQASSPTNNIFASSK MYLGEIPROV MYLGEIVRNITHQASSPTNNIFASSK MYLGEIPROV MYLGEIVRNITHQASSPTN MYLGEIPROV MYLGEIPROV MYLGEIPROV MYLGEIPROV MYLGEIPROV MYLGEIPROV MYLGEIPROV MYLGEIPROV MYLGEIPROV MYLGEIPROV MYLGEIPROV MYLGEIPROV MYLGEIPROV MYLGEIPROV MYLG	111	1461	-	1/26	2	344	
TSSSNPYATT VFGQSSDPVSSYGFVNTAESST	1 ***	1401	A	1420	1 -	344	
112			ĺ	İ		ļ	
112			[
SRRSSQRTCSRACSGAWSRTW*RSS*TSSSSC STSCSSSSSSCGPGGPLGARGVHTSCLNSC STSCSSSSSSCGPGGPLGARGVHTSCLNSC MSSSTTSSTTSTT	112	1462	Ι	1/3/	46	372	
STSCSSSSRSGRPGGPLGARGVHITSCLNSC MSSTTSSTTSTT	112	1402	\ ^	1454	1 70	372	
MSSSTTSSTTSTF							
113	}	}	ł	}	}		
MYLGEIVRNILIDFTKKGFLIRGQISEMLKTR	113	1463	A	1439	3	292	
114	113	1100	11	1437	1		
114	1		1				
QRTKVH_PGHKTGPAVAKDTPEPVÄKEFTVP	114	1464	A	1463	1	396	
115	***	1		1	⁻		
EDP*KNA*LKQMHAATTHWQQHQQHQVGC QYHGIMQ QYHGIMQ QYHGIMQ AGSYPSMVWSCHWGVTQKRRAL*VYSFEEG GRRKCQQYWPLEKDSRIRFGFLTVSNLGVEN MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN NYCN NYCN LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ YWTKYQVWEWLQHFLDTNQLDANCFPGEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC HLKWNGDSLFLCLSLPC HLKWNGDSLFLCLSLPC HLKWNGDSLFLCLSLPC QFSNSFQVPLQAKLVSSHKFQQNQKHK QLQATSVFPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK QHWIEGHTCLDNNIHQAASEPINNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H GHTSPHYSHAMATRIQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL RTSPLYDREV WERGHTCLDNNSGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLDAGASRPKYFYP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFETISTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMSPTDQQVHCWAWLKKHMFKDSN	1	1	1	-	{		ATSOGP*SPFSEEPPLPPSNEEVPPTLPP*EPQS
115	İ	ł	l	1	i	1	
GRRKCGQYWPLEKDSRIRFGFLTVSNLGVEN MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN NYCN 116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ YWTKYQVWEWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC 117 1467 A 1479 I 381 GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H 119 1469 A 1486 I 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSFLYDRLDAQGARWMEKHGFERRKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVUDMSSFTIEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKHMPKDSN			1		ļ		QYHGIMQ
MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN NYCN	115	1465	A	1464	291	2	
NYCN			1				
116]		j	1	
TYNTKYQVWEWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC 117 1467 A 1479 1 381 GTSGGPKRVLVTERPPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEELASDPNNEESL*RPWALEDFEIGRPLG KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H 119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTTFKPDWFDIVESEVKCCK EAVCUMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN			<u> </u>				
DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC 117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLQFSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H 119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSLARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	116	1466	A	1465	667	337	LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ
HLKWNGDSLFLCLSLPC		Ì	Í		Ĭ		YWTKYQVWEWLQHFLDTNQLDANCIPFQEF
117	1	1		İ			
VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H 119 1469 A 1486 I 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN		<u> </u>	<u> </u>		<u> </u>	·	
QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN	117	1467	A	1479	1	381	GTSGGPKRVLVTERFPWQNPLPVNRGQAQR
NPEELASDPNNEESL*RPWALEDFEIGRPLG KGK			1				VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK
Items							
118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H 119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN					Į.		
PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H 119 1469 A 1486 I 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	1.5	1.65	 	1 105	 	205	
QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H 119 1469 A 1486 I 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	118	1408	A	1485	5	383	
NLAFLATGVVRHMRKLFMGANLEGPGPTVS H		[1			
H	1	l l					
119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	1	1	1	1		Ì	
NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	110	1440	+	1406	 	308	l
KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	119	1409	A	1400	} *	270	
PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	1		1	1	}		
DDEDLLMYL 120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	ì	1			ļ		
120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN]]	1	l		
LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	120	1470	ΙΔ-	1407	+3	999	
FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	1.20	17/0	^	1 17/	1	1	
RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	1	1	(į ·		
PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN		1		1	1		
EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	-	1	1	1	1		
NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	1	1		1	1		
NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	1	1					
		}	1	1	1	1	
		1				1	

SEQ ID	SEQ ID	Met	LCEO	Predicted	Predicted end	
NO: of	NO: of	hod	SEQ ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		•		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ł	İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ł	ĺ	Ì	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						MTPDHFPSLFCKEMSVGYANGIRVMSMTHT
		1				GEPGFMLYIPIEYRWGFTMLSTLVSNS
121	1471	A	1498	3	306	AQFLLVGWDHIL*LIVL*TNLTELGRTTCDQN
						WPNSPDVLNHGCFYMQCLSKDCTIGYVSRE
	[ĺ				MLVAHTHTVEEHTGTHLQYVSWPDHSVPDD
			1			SSDFVEFEN
122	1472	A	1533	121	329	LGLFSFVWTEVLEEPKDFSCETEDFKTLHCT
	İ	ł				WDPGTDTALGWSKQPSQSYTLFES*VGSGYII
					ļ	DNFFLA
123	1473	A	1547	111	408	DARTTWKPRNGSSGIWPGDGAK*PPAVEQAE
			l			RGHVEMIEKLTFLNLHTSEKDKGGNTALHLA
						AKHGHSPAVQVLLAQWQDINEMNEKQQTPL
· .						HVAADRG
124	1474	A	1555	1	745	MTFDDDDKNTYGVALVWKKFQTQSLRLSDL
ļ		}				HRKSHLWRGIVSITLIEGRDLKAMDSNGLSDP
1		ļ				YVKFRLGHQKYKSKIMPKTLNPQWREQFDF
						HLYEERGGVIDITAWDKDAGKRDDFIGRCQV
						DLSALSREQTHKLELQLEEGEGHLVLLVTLT
						ASATVSISDLSVNSLEDQKEREEILKRYSPLRI
						FHNLKDVGFLQVKVIRAEGLMAADVTGKSD
						PFCVVELNNDRLLTHTVYKNLNPEWNKVFTL
		İ				*VALVWKKFQTQSLRLSDLHRKSHLWRGIVS
						ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY
						KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT
						AWDKDAGKRDDFIGRCQVDLSALSREQTHK
			1			LELQLEEGEGHLVLLVTLTASATVSISDLSVN
				,		SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV
						KVIRAEGLMAADVTGKSDPFCVVELNNDRLL
125	1475	<u> </u>	1.55		700	THTVYKNLNPEWNKVFTL
125	14/3	A	1556	57	509	GGPAPNSRYAEP*KNSLAMT*AHADCENYVA
						CGGLDNICSIYNLKTREGNVRVSRELPGHTGY
						LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT
						TTFTGHSGDVMSLSLSPDMRTFVSGACDASS
126	1476	A	1592	3	178	KLWDIRDGMCRQSFTGHVSDINAVS
120	14/0	Λ.	1372	3	1/0	KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL
127	1477	Δ	1612	1	107	EMLPTCDLADQHNIKFHYAFALNR*ER
141	14//	A	1012	1	497	TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS
						VLGAYISFGVPSSHLLTASVMSAPASLAAAKL
*						FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA
						WVGNMFDYPQLSFELICSYIFMPFSFMMGVE
						WPDSFM
128	1478	A	1619	286	486	CCMNSKAQESVFKNVLCNPPALSEMPDVKA
	1473	1.	1017	200	700	EDEVDFRASSISEEVAVGSIAATLKMKOGPM
						TQAINR
129	1479	Α	1627	1	395	PTRGALRYWIFGRFLCNIWAAVDVRCCTATI
			/	. *	373	MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA
						LLCVWALSLVIYIGPLLGWRHPAPEDETICQI
				İ		NEEPGYVLFSTPGSFYLPLAIMLVMN*RVYRV
				İ	(AKTE
130	1480	Ä	1638	2	466	DPRVRTKIVNRKTTIYEIQDKTGSMAVVGKG
				-	.50	ECHNIPCEKGDKLRLFCFRLRKRENMSKLMS
					Ì	EMHSFIQIQKNTNQRSHDSRSMALPQEQSQHP
						KPSEASTTLPESHLKTPQMPPTTPSSSSFTKVT
					ļ .	KDKDIK*LLFNLYSSVEILPEVLHLKT
131	1481	A	1651	607	3	LAEGGDVFDCVLNGGPLPESRAKALFROMVE
				- - · ·	-	AIRYCHGCGVAHRDLKCENALLQGFNLKLTD
					1	FGFAKVLPKSHRELSQTFCGSTAYAAPEVLQ
						GIPHDSKKGDVWSMGVVLYVMLCASLPFDD
					·	TTO TO THE TOTAL TO THE TO

					18 0	(1.4)
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ľ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Ĭ	1	[ĺ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
Į.	ŀ			peptide	Doguesion	/=possible nucleotide deletion, \=possible
1	}	1	ł	sequence	ł	nucleotide insertion
	<u> </u>	<u> </u>		sequence		TDIPKMLWQQKGVSFPTHLSISADCQDLLK
J		j	l	ļ	j	
ĺ	1					RLLEPDMILRPSIEEVSWHPWLAST**KQWQV
			<u> </u>	<u> </u>		LSNKVGGESKPKKKK
132	1482	Α	1656	150	48	LVAKSLLYCGCLFFLLQLAKNVGNNSFNDIM
İ						EANLTSPSPKPTPSSDM*VFLIY*TYFGAWHV
i		1	ł	l		VDAQ
133	1483	A	1660	3	406	RKHIKLLIQKLSDVP*ECQNNQL*KLTEICEKE
1]	1	1		***	KKEFKKKMDDQRPEKITEA*SKDKSPMEEEK
1	1		1			TEMIRSYIOEVGRYIKRLEEAQSKRLEKLREK
	1	1	1			HKEIRQPILDEKPKGEGSSSFLSETCHEDTSWF
ì	1	ľ	ì	t	ĺ	
	1	<u> </u>		1076		PNFTP
134	1484	Α	1666	1276	466	PGSTHASARITIY*L*IILSNATEVDNNFSKPPP
1	1	l .		1	1	FFPAGAPPASSSSSSSSSSPPTVSTAPPLIPPPGF
1		ļ		1		PPPPGAPPPSLIPTIESGHSSGYDSRSARAFPYG
{	-	1		1	[NVAFPHLPGSAPSWPSLVDTSKQWDYYARSS
	1	1				SSSSSSSSSSSSPRDRDRER*RTRERERERDHS
	Į.	}	1	}	l	PTPSVFNSDEERYRYREYAERGYERHRASRE
	ţ	i	ŀ			KEERHRERRHREKEETRHKSSRSNSRRRHESE
	1	1	1	-		EGDSHRRHKHKKSKRSKEGKEAGSEPAPEQE
	1		1			STEATPAE
126	1405	<u> </u>	1.555		415	
135	1485	Α	1673	1	417	PTRPVNSSQAFALVYYTLGALGGNLIAHMGL
ł	1	Ì	ŀ			GYRYWAGIGVLQSCESALTHYRLVANHVAS
	1	İ	Į.			DISLTGGSVVQRIRLPDEVENPGMNSGMLQE
1	1	1	ì		ļ	DLIOYYQFLAEKGDVQAQVGLGQLHLHGGR
	i	1	i		Ì	GV*QNHQRAFDYFNLAA
136	1486	A	1678	525	9	ANTSLSSAAVSAVSPPPCRTSTATTLPPPMPSF
ł	1	l		ł	}	FCVFPSPSMSPSPSEFLSCIASVSRVHSLSSSSS
i		1	ľ	Į.	ľ	GSSSTASSLNFSAIMGSSSATASWVLSTASTPP
	1	İ		İ		CPSALPSSPAQES*SLAASSSAWPVAGISPSGA
	1	I	{			CTFPAGSASGAAKAPSPSWRCPSFRALFSLLD
i	1	1	l .		1	SSSLSL
137	1487	A	1680	1	2999	AHRDEIQRKFDALRNSCTVITDLEEQLNQLTE
13/	146/	A	1000	1 .	2333	DNAELNNONFYLSKOLDEASGANDEIVOLRS
1]	}	}	· .	,	
	ł	ı				EVDHLRREITEREMQLTSQKQTMEALKTTCT
	1				1	MLEEQVMDLEALNDELLEKERQWEAWRSVL
	1	1	1 '	i	1	GDEKSQFECRVRELQRMLDTEKQSRARADQ
1	1	1	1	1	ļ	RITESRQVVELAVKEHKAEILALQQALKEQK
ŀ	ł	1	1	l	1	LKAESLSDKLNDLEKKHAMLEMNARSLQQK
	1	1	1	1	1.	LETERELKQRLLEEQAKLQQQMDLQKNHIFR
	}		ì		1	LTQGLQEALDRADLLKTERSDLEYQLENIQV
	1					LYSHEKVKMEGTISQQTKLIDFLQAKMDQPA
1		l		1	1	KKKKVPLQYNELKLALEKEKARCAELEEALQ
	1	1	1	1		KTRIELRSAREEAAHRKATDHPHPSTPATARO
		1	1		1.	OIAMSAIVRSPEHOPSAMSLLAPPSSRRKESST
	1	1	1	1	}	
1		l			1	PEEFSRRLKERMHHNIPHRFNVGLNMRATKC
		1	1			AVCLDTVHFGRQASKCLECQVMCHPKCSTC
1	1	1	1	1	1	LPATCGLPAEYVTHFTEAFCRDKMNSPGLQT
1		1		1	1	KEPSSSLHLEGWMKVPRNNKRGQQGWDRK
1	1		1		1	YIVLEGSKVLIYDNEAREAGQRPVEEFELCLP
		1	1			DGDVSIHGAVGASELANTAKADVPYILKMES
1	1	J	1	j	}	HPHTTCWPGRTLYLLAPSFPDKQRWVTALES
		}	ł		1	VVAGGRVSREKAEADAKLLGNSLLKLEGDD
		1	1		1	RLDMNCTLPFSDQVVLVGTEEGLYALNVLK
ł	1	I	1	I	ì	NSLTHVPGIGAVFQIYIKDLEKLLMIAGEERA
	Ì	1	!	İ		
			1	j		LCLVDVKKVKQSLAQSHLPAQPDISPNIFEAV
		1	1			KGCHLFGAGKIENGLCICAAMPSKVVILRYN
].	,	1	1	1		ENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNK
1	ì	1	1	1	1	FYEIDMKQYTLEEFLDKNDHSLAPAVFAASS
L	l	l	L	1		NSFPVSIVQVNSAGQREEYLLCFHEFGVFVDS

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	L Amino said converse (AmAlonina Ca-Custalina
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	}	}		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
	_	}		sequence		nucleotide insertion
		Ĭ				YGRRSRTDDLKWSRLPLAFAYREPYLFVTHF
		ļ		1		NSLEVIEIQARSSAGTPARAYLDIPNPRYLGPA
		j		!	}	ISSGAIYLASSYQDKLRVICCKGNLVKESGTE
				<u> </u>		HHRGPSTSRR*PASPLPQYQGQRAFLQGRRK
138	1488	A	1686	2	526	GRPQGPAPGAGSPPESGPGLWAALGCSLVWV
					{	PLCCLGGAAGRL*ARSGKSGLRRRRAHAGPP
	ļ	<u> </u>	1			PGGPCNSCP*CSAPESGGRGPLPGPGTGGVCS
	ļ		İ			CWTRGCQTTARTAAAAAAPGPAGRRPPGGA
]	ļ]	PQNGSCAASASQEAAAPPPMCPPGRRWAVAS
120	1400	ļ.,	1602		0.00	PPETRCPAAPGTRCRRLEAA
139	1489	A	1693	3	376	LPSMSNCTSCFRLQSRTES*IRQAGHLLGRNE
			1			FIETKALGCAWFSLCYYLVLYFESSHKVDFVF
				•		IV*CFSTPPGAQMTIMSQACAERCNIMRLVDR
140	1490	A	1704	3	376	RWAGIAKGVGTQKIIGRVHLGEQKALGL
140	1450	^	1704	3	1 3 / 6	ERTNKFIKELIMDGKNLIAATKSLSVAQRKFA HSLRDFKFEFIGDAVTDDERCIDASLREFSNFL
		ì				KNLEEQREIMVS*EGCKLISQLSRGKKIWIWK
	[ĺ		[LVLVEVVKHLSLGTVVHCNGKMRFPEP
141	1491	A	1743	1	362	LITNKVFVARELSCLDVHLDSTGSTAVVADQ
• • •	1171	'	1,43	1 -	302	DKLELELVLKGSYEDTQTSFLGTASAFRFHY
	<u>.</u>					MAAL*TELSGRLRSSKSNGWNGDNSTGYLTV
	ĺ	l	İ			PLRPLTIVKEVTMDVPAPNVRGLNWMG
142	1492	A	1769	1	406	NNPSTLPRGS*PMSPRTTMGRRRQRRREHKSS
	ĺ	(1	ĺ	LSLASSTVGPGGQIVHTETTEVVLCGDPLSGF
]			ļ	GLQLQGGIFATETLSSPPLVCFIEPDSPAERCG
•			Ì			LLQVGDRVLSINGIATEDGTMEEANQLLRDA
	L		l	<u>. </u>	}	ALAHKVV
143	1493	A	1789	1	447	QMLRNGGDQNTVPDYHFADRIRELL*PTEDQ
	}					KNCIP*DTYLRPSALGNIVEEVTHPCSPGPCPA
						NELCEVNRKGCTSGDPCLPYFCVQGCKLGQA
		1				SDFIARQGTLIQVPSSAGEVECYKICSCGQSGL
		<u> </u>				LENCMEMHCMDLPTDTSALVR
144	1494	A	1814	1	404	PGRRFRPRLSQAGTDSGS*VFPDSFPSAPAEPL
		ļ				PYFLQEPQDAYIVKNKPVELRCRAFPATQIYF
	1		i :			KCNGEWVSQNDHVTQEGLDEATGLRVREVH
		1				IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK
145	1495	A	1827	26	448	SRRAYVRI
140	1473	^	102/	20	448	XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N
						CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE
		}	}			THLALCPIVQHPEDTCIHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD
		1				PEDARVLCRQLNCGTAL
146	1496	A	1828	574	333	QHEGGDLRRRQLGEIQLTVRYVCLRAASAC*
	1	1	1020	J. 4	555	SMAAET*HHVPASGADPYVRVYLLPERKWA
		1				CRKKTSVKRKTLEPLFDET
147	1497	A	1855	1	372	ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH
	,	l	.555	1	3/2	VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE
		1	[TSVTYSMG*HGAPTGSEAGANWNH**LHAH
		·	j			YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE
						Q
148	1498	A	1879	568	7	LLSALDDKGGTQPSASFSNAPTIVCVTACPAG
]]			IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN
	}					GIEGRLTADQLNSATACIFAAEVAIKESERFN
)	}			GIPALSVPVAEPIRHAEALMQQALTLKRSDET
		l				RTVQQDTQPVKSVKTELKQALLSGISFAVPLI
·						VAGGTQVA*AV*RQGISSLHDVQVRTWNS
149	1499	A	1880	611	24	GLNSENALSNEAMERGWQCLRLFAERLQDIP
		i .	1		I	DOORDAWATATI DI ADDIA ODELAY AORU COD
	1	ļ				PSQIRVVATATLRLAVNAGDFIAKAQEILGCP VQVISGEEEARLIYQGVAHTTGGADQRLVVD

NO. of No. of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nuclectide coation correspondit coation correspondit coation correspondit coation correspondit coation correspondit coation coatio					1		
Sequence							
Sequence			١.				
150 1500 A 1894 2 750 1700 1	1 :		,				
amino acid of peptide residue of peptide residue of peptide sequence T-Threonine, V=Valine, W=Typtophan, Y=Typtophan, Y=Dipo codon, V=possible mulcotide deletion, V=possible nucleotide d							
Persidue of peptide sequence							
peptide	j j			}			Y=Tyrosine X=Unknown *=Ston codon
						- Squemoo	
150]				,		
150	 		 	<u> </u>	304401100		
150		,					
150				}	}		, , , ,
LMPPPFLVDVDGNPHPTKYQRLVPGRENSAG EHLPDGGVVATSDEGVEQIBLICTIONDERS PESSIL DGMRQLQQQDDRMGADQDTIPRG LSNGEETPRRGFRRLSLDIQSPPNIGLRRSGQV EOVRQMHQLQQQDDRMGADQDTIPRG LSNGEETPRRGFRRLSLDIQSPPNIGLRRSGQV EOVRQMHQLAPRSQIATERDLQAWKRVVV PEVPLGIFRKI.EDFRLEKGEERENI,VIIGKRK TLQLSKKSDSVQLVSQSRPRTCRKYP EVPLGIFRKI.EDFRLEKGEERENI,VIIGKRK TLQLSKKSDSVQLVSQSRPRTCRKYP EVPLGIFRKI.EDFRLEKGEERENI,VIIGKRK TLQLSKKSDSVQLVSQSRPRTCRKYP EVPLGIFRKI.EDFRLEKGEERENI,VIIGKRK TLQLSKKSDSVQLVSQSRPRTCRKYP EVPLGIFRKI.EDFRLEKGEERENI,VIIGKRK TKNEHKTILKTILKDEQTRKLAIL.ABGYYEGGV TFKNEHKTILKTILKDEQTRKLAIL.ABGYYEGGV NEMMASQALRIDEQGEACQALRQLQQGM ELDNAYQSKIKMQTEAQHERELQKLEQRXYS RRAHLEQKIEELAALQKERSERIKNLLERQE RRIETPDMSSLRMGFGRIUTI.DFFKEDYR ELDNAYQSKIKMQTEAQHERELQKLEQRXSGBXL RRAHLEQKIEELAALQKERSERIKNLLERQE RRIETPDMSSLRMGFGRIUTI.DFFKEDYR ELDNAYQSKIKMQTEAQHERELQKLEQRXSDBXL RRAHLEQKIEELAALQKERSERIKNLLERQE ELDDEVQNAAAGALAMLTAAHKKLCLKMT QVIT QVIT EVPLEY QVIT EVPLEY QVIT ELTPFQLYTPELIFKHFQURRLITNFLFFVPFG FNFLLYMIFLYT ELTPFQLYTPELIFKHFQURRLITNFLFFVPFG FNFLLYMIFLYT TYPEVEGGSLNGKQRYSKMTGAYLGEIV VYTEVEGGSLNGKQRYSKMTGAYLGEIV VYTEVEGGSLNGKQRYSKMTGAYLGEIV VYTEVEGGSLNGKQRYSKMTGAYLGEIV ROLLDILTSGLLFRGAYLGEIV ROLLDILTSGLLFRGAYLGEIV ROLLDILTSGLLFRGAYLGEIV ROLLDILTSGLLFRGAYLGEIV ROLLDILTSGLLFRGAYLGEIV RECHARCA	150	1500	A	1894	2	750	
	130	1500	Α	10/4	_	130	
PESSILDGMIRQLQQQQDQMGADQDTIPRG LSNGGETPRRGFRISLSDIOSPHIGLRSGQQV EGYRQMHQNAPRSQLATERDLQAWKRRVVV PEVPLGIFRKLEDFRLEKGEERNILYTIGRKRK TLQLSHKSDSVGLVSQSPRTCRRKYP TLQLSHKSDSVGLVSQSPRTCRRKYP TLQLSHKSDSVGLVSQSPRTCRRKYP TLQLSHKSDSVGLVSQSPRTCRRKYP TRNSHKTILKTLKDEQTRKLAILARQYRQSI NEMMASQALRDEAQEAGCQALRLQLQQEM RELINAYQSKIKMQTEAQHERELQKLEQRVSL RAHLEQKIEELAALQKERSERIKNLLERQE REIETFDMESLRMGFGONLVTLDFFKEDYR LVRLLDTORDGLONYEALLGTLSGRSDKL RAHLEQKIEELAALQKERSERIKNLLERQE REIETFDMESLRMGFGONLVTLDFFKEDYR LVRLLDTORDGLONYEALLGTLSGRSDKL ROKIFKERALPDIENYMFENHDQLRQAATEC MCNMVLHKEVQEFFLADGNDEALACLLCKMT QVTT AVGSLRLEYLQIPPVSRAYTTACVLTSAAVQL ELITPFQLVFFPLIFRHFQIWRLITHFLFFVPG FNFLLYMIFLY TEVELOFFT TE					· ·	,	
LSNGETPRRGPRISIDIOSPPNIGLRSGQV EGYRQMHQNAPSQIQATERDLQAWKRRVVV PEVPLGIFKL EDFEL EKGEERNI, YIIGKKRK TLQLSHKSDSVGLVSQSRPRTCRRKYV PEVPLGIFKL EDFEL EKGEERNI, YIIGKRKK TLQLSHKSDSVGLVSQSRPRTCRRKYV PEVPLGIFKL ENGERENNI, YIIGKKRK TLQLSHKSDSVGLVSQSRPRTCRRKYV PEVRLSHVIKLKDEQTRISLABQYPGSI NEMMASQALRIDEAQEAECQALRLQQEM ELLNAYQSKIKMQTEAQHERELQKLEQRVSL RRAHLEQKIEEELAALQKERSERINLLERQE REIETPOMESLRMGFOINLYTHLERQE REIETPOMESLRMGFOINLYTHLERQE REIETPOMESLRMGFOINLYTHLERQE REIETPOMESLRMGFOINLYTHLERQE ROUNT CANNOWNLHKEVQERFLADIONBRIKL VULLCG EDDDKVQNAAAGALAMLTAAHKKLCLKMT QVIT CANNOWNLHKEVQERFLADIONBRIKL VULLCG EDDDKVQNAAAGALAMLTAAHKKLCLKMT QVIT AYQSLRLEYLQIPFVSRAYITACVLTSAAVQL ELITPFOLVTPELIFKHFQWRHJITINFLEFYPFG FPFLLYMFLYT FYPTE FPFLLYMFLYT YDTEVDEGSLNPGKQRYEKMTSGMYLGEIV RQILIDITKQGLJRRGQISERLRTRGFFSTKFLS QEEDRIALJQVRRILQQIADL GEDRIALJQVRRILQQIADL GEDRIALJQVRRILQQIADL GEDRIALJQVRRILQQIADL GEDRIALJQVRRILQQIADL GEDRIALJQVRRILQQIADL GEDRIALJQVRRILQQIADL GEDRIALQVRRILQQIADL GEDRIALQVRRILQQIADL GEDRIALQVRRILQQIADL GEDRIALQVRRILQQIADL GEDRIALQVRRILQQIADL GEDRIALQVRRILQQIADL GEDRIALQVRRILQQIADL GEDRIALQVRRILQQIADL GEDRIALQVRRILQQIADL GEDRIALQVRRILQQIADL GESTALVRILGHT GETRIALGHT	}						
EGYROMHONAPRSOIATEDLOAWKRRYVV PEVPLGIFKILEDPRILEKGEERNI/UIGRKRK T.Q.I.SHKSDSVGLVSOSRPRTCRRKYP FEVPLGIFKILEDPRILEKGEERNI/UIGRKRK T.Q.I.SHKSDSVGLVSOSRPRTCRRKYP T.Q.I.SHKSDSVGLVSOSRPRTCRRKYP ST. GKTTQIQTTMONEYKTVQKQYKTIFKNKKA MEMQIKKQFQDTCKVQTKQYKALKNHQLEV TYKNEHKTILKTILKDEQTRKLALLAQVEQSI NEMMASOALRDEAQEAECQALRLQLQCEM ELLNAYQSKIKMQTEAQHERELQKLEQRVSL RRAHLEQKIEELAALQKERSERIKNILERQE REIETFDMESLRMGFGNLVTLDFPKEDYR ELLNAYQSKIKMQTEAQHERELQKLEQRVSL RRAHLEQKIEELAALQKERSERIKNILERQE REIETFDMESLRMGFGNLVTLDFPKEDYR RRAHLEQKIEELEAALQKERSERIKNILERQE REIETFDMESLRMGFGNLVTLDFPKEDYR ROKIFKERALPDIENYMFENHDQLRQAATEC MCMNYLHKEVQERFLADDINEKLVVLLCG EDDDKVQNAAAGALAMLTAAHKKLCLKMT QVTT ATQSLRLEYLQIPPVSRAYTTACVLTSAAVQL ELITPFQLYFIPLIFKHFQIWRLITHNILFFVPG FNFLLYMIFLY TEVEDVSSLNSHORQKQYEKMKLIVLLCG EDDDKVQNAAAGALAMLTAAHKKLCLKMT QVTT ATQSLRLEYLQIPPVSRAYTTACVLTSAAVQL ELITPFQLYFIPLIFKHFQIWRLITHNILFFVPG FNFLLYMIFLY TEVEDVSSLNSHORQKQYEKMTSGMYLGEIV ROQLIDITRQGLIPRGQISERLRTRGFFTKFLS QESBRLALLQVRRILQVARILAQRILQVARILLUCAQRADILAQRILAQRILQVARILUCAQRATILQVARILDVARILQVARILQVARILQVARILQVARILQVARILQVARILQVARILQVARILQVARILAQRARINILQSLKFQGG HQPAATLLR ALAQGRHAQTVETLLRIGAHINILQSLKFQGG HQPAATLLR ALAQGRHAQTVETLLRIGAHINILQSLKFQGG HQPAATLLR ALAQGRHAQTVETLLRIGAHINILQSLKFQGG HQPAATLLR KFLKDLEQVARILGVARILQVARILAQRARILAVARILLQVARILAQRARINAVACLINER KKKLESIDKL LUVETLQQQMIDIAQGRAYAVSFILLYSSPDVKE LUVETLQQQRISHGTGVCKCYPJQGSAALIMKLR KKKLESIDKL LUVETLQQQRISHTGVCKCYPJQGSAALIMKLR LUVETLQQQRISHTGVTGCKCYPJQGSAALIMKLR LUVETLQQQRISHTGVTGCKYPJQGSAALIMKLR LUVETLQQQ							
	1						
151							
151]				l		
MEMQIKKOPGDTICKVQTKQVKALKNHQLEV TPKNPIHKTILKTLKDEGTKALALAGQYQSI NEMMASQALRLDEAQEAECQALRLQLQQEM ELLNAYQSKIKMQTFAQHERELGKLEGKYSL RRAHLEQKIEELAALQKERSERIKNLLERQE REIETFDMESLRMGFGNLVTLDFPKEDYR ELLNAYQSKIKMQTFAQHERELGKLEGKYSL RRAHLEQKIEELAALQKERSERIKNLLERQE REIETFDMESLRMGFGNLVTLDFPKEDYR TVRLLDTORGIG,NYPELJGLTNLSGRSDKL RQKIFKERALPDIENYMFENHDQLRQAATEC MCNMVLHKEVQERFLADGNDYLKLVVLLCG EDDDKVQNAAAGALAMLTAAHKKLCLKMT QVTI 153 1503 A 1921 1 237 AYQSLRLEYLQIPPVSRAYTTACVLTSAAVQL ELITPFQLYFIPELIFKHFQIWRLITNFLFFVPFG FNFLLYMIFLYT TYDTEVDEGSLNPGKQRYEKMTSGMYLGEIV RQLIDLTKQGLLFRQGSERLRTRGIFETKFLS QESDRLALLQVRRILQQLGLD TYDTEVDEGSLNPGKQRYEKMTSGMYLGEIV RQLIDLTKQGLLFRQGSERLRTRGIFETKFLS QESDRLALLQVRRILQQLGLD TELAKIKMEAKKYEKELETHFQNDFEKACQA KSEALVLREKSTLERHKHQBIETKEIYAQRQ LILKDMDLLRGREAELKGRAFESYQLELK DDYURTYRLIEDDRNIQISGHWQESP TELAKIKMEAKKYEKELTHFQNDFEKACQA KSEALVLREKSTLERHKHQBIETKEIYAQRQ LILKDMDLLRGREAELKGRAFESYQLELK DDYURTYRLIEDDRNIQISGHWQESP TELAKIKMEAKKYEKELTHFQNDFFEKACQA RSCGCLIWFFITTDLQILTSSILPSIL TELAKIKMEAFTIRKLHPTDNFAQ RSCGCLIWFTPTTDLQILTSSILPSIL TELAKIKMEAFTIRKLHPTDNFAQ RSCGCLIWFTPTTDLQILTSSILPSIL TELAKIKMEAGOPGKSSEPOKATQUDFTA AYEYDAGNIWCKDCNTICGTMFDFFTHMH NKHTIQQQFOKSSEPOKACQDTFLPERQO THRILNYTAEPPCTSMPIYWMPDVPHRCTTA NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED RHCQTQACPPLSWPRQTLDILGTARAIQFLH QDSPSLHIGDIKSSNVLLDERLTPKLGDFGLA RPSRFAGSSPIOSSM RTRILNYTAEPPCTSMPIYWMPDVPHRCTTA NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED RHCQTQACPPLSWPRQTLDILGTARAIQFLH QASPSLHIGDIKSSNVLLDERLTPKLGDFGLA RPSRFAGSSPIOSSM THRILNYTAEPPCTSMPIYWMPDVPHRCTTA NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED RHCQTQACPPLSWPRQTLDILGTARAIQFLH QASPSLHIGDIKSSNVLLDERLTPKLGDFGLA RPSRFAGSSPIOSSM RETISTIVE PROGRATILH AAAHGHSEVVEELVSADVDLFFDEQGISALH LAAQGRHAQTVETLRHGAHINLQSLKFQGG HGPAATLL AAAHGHSEVVEELVSADVLLFDEQGISALH LAAQGRHAQTVETLRHGAHINLQSLKFQGG HGPAATLA RETGSVSLSPGGLEGAESYAVSPILYSSPDVKE LWETTQGMBASHTGYCKSTPQGSAAILMKLR URUTTQHADYSNIKAAYEAMKNRACLINER KRILESIBKIA RETG	151	1501	Á	1900	141	785	
TPKNEHKTILKTLRDEQTRKLAILAEQYEQGS NEMMASQALRIDEAGEQAJALIQLOGEM NEMMASQALRIDEAGEQAJALIQLOGEM NEMMASQALRIDEAGEQAJALIQLOGEM ELLNAYQSKIKMQTEAQHERELQKLEQRVSL RRAHLEQKIEGELAALQKERSERIKNILERQE REIETTDMESLRMGFGONLVTLDFPKEDYR	121	1501	Λ	1700	171	601	
		'		•		ļ	LDKNEHKLII KLI KDEULDKI YII YEUNEUGI
RRAHLEQKIEFELAALQKERSERKINLIERQE REIETPOMESLRMGGNLVTLDFFKEDYR 1502 A 1915 2 377 LVRLLDTQRDGLQNYEALIGLTNLSGRSDKL RQKIFKERALPDIENYMFENHOQLRQAATEC MCNMVLHKEVQERLAGONDRILKUVILLGG EDDDK VQNAAAGALAMLTAAHKKLCLKMT QVTT QVTT AYQSLRLEYLQIPPVSRAYTTACVLTSAAVQL ELITPFQLYFIPELIFKHFQIWRLITINFLFFVPFG FNFLLYMFLYT FNFLLYMFLYT YDTEVDEGSLNPGKQRYEKMTSGMYLGEIV RQLIIDLTKQGLLFRGQISERLRTRGIFETKFLS QIESDRLALLQVRRILQQLGLD TEAKIKMEAKKKVKELTMFQNDFFEKACQA KSEALVLREKSTLERIHKHQEIETKEIYAQRQ LLIKDMDLLRGREAELKQR VEAFESYQLEIK DDYIIRTYRLIEDDRINQISCHWQESP 1506 A 1935 1 270 VTRKLPIFIVDAFTARAFRGSPADCLLENEL DEDMHQKIAREMALSTAFRKLIPTDNFAQ RSCFGLIWFTPTDLQILTSSILPSIL SEKVNNEKFRTKSFRPAESPGATKQLDQPTA AYEYYDAGNHWKODONTICGTMFDFFTHMH NKKHTQGGPGKSSDFQKEELQQTFLPFERQG TTRIKLNTAFPCTSMPJYWMPDVPHRCTTA ATGRICA ATGRI	l i						
REIETPDMESLRMGFGNLVTLDFFKEDVR	1						
1502	1						
RQKIFKERALPDIENYMFENHOQLRQAATEC MCNMVLHKEVQERFLADGNDRLKLVVLLCG EDDDKVQNAAAGALAMLTAAHKKLCLKMT QVTT QVTT QVTT AYOSIRLEYLQIPPVSRAYTTACVLTSAAVQL ELITPFQLYFIPELIFKHFQIWRLITNFLFFVPFG FNFLLYMIFLYT FNFLYMIFLYT F	152	1502	Δ	1015	2	377	
MCNMVLHKEVQERFLADGNDRİKLÜVÜLLG EDDDKVQNAAAGALAMLTAAHKKLCLKMT QVTT	132	1502	Λ.	1713	4	311	
	1					1	
153							
1503	} }						OVTT
	153	1502	Δ	1021	1	227	
1504	123	1505	Λ.	1741	*	102	
1504	ļ i]			
YDTEVDEGSLNPGKQRYEKMTSGMYLGEIV	154	1504	Δ	1978	2	354	<u></u>
RQILIDLTKQGLLFRGQISERLRTRGIFETKFLS QIESDRLALLQVRRILQQLGLD	'~'	1004	Δ .	1720		334	
OIESDRLALLQVRRILQQLGLD		.					BUILINI TRUCH I EBCUIGEDI DEDCIEEEREN 6
155 1505 A 1929 2 369 TELAKIKMEAKKYEKELTIMFQNDFEKACQA KSEALVLREKSTLERIHKHQEIETKEIYAQRQ LLLKDMDLLRGREAELKQR VEAFESYQLELK DDYIRTYRLIEDDRNIQISGH WQESP 156 1506 A 1935 1 270 VTRKLPIFIVDAFTARAFRGSPAADCLLENEL DEDMHQKLAREMNLSETAFIRKLHPTDNFAQ RSCFGLIWFTPTIDLQILTSSILPSIL 157 1507 A 1936 584 305 ESKVNNEKFRTKSPKPAESPQSATKQLDQPTA AYEYYDAGNHWCKDCNTICGTIMFDFFTHMH NKKHTQGQFQKSSDFQKEELQQTFLPPERQG 158 1508 A 1939 1 423 TTHRLNVTAEPPCTSMPYWMPDVPHRCTTA NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED RLHCQTQACPPLSWPQRLDILLGTARAIQFLH QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA RFSRFAGSSPIQSSM 159 1509 A 1974 3 401 HTSTARLLHRGAGKEAVTSDGYTALHLAAR NGHLATVKLLVEEKADVLARGPLNQTALHL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR 160 1510 A 1982 2 417 KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA TYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRISHTGVKSTPGQSAAILMKLR							
KSEALVLREKSTLERIHKHQEIETKEIYAQRQ LLLKDMDLLRGREAELKQRVEAFESYQLELK DDYIRTYRLIEDDRINIQISGHWQESP 156 1506 A 1935 1 270 VTRKLPIFIVDAFTARAFRGSPAADCLLENEL DEDMHQKIAREMNLSETAFIRKLHPTDNFAQ RSCFGLIWFTPTTDLQILTSSILPSIL 157 1507 A 1936 584 305 ESKVNNEKFRTKSPKPAESPQSATKQLDQPTA AYEYYDAGNHWCKDCNTICGTMFDFFTHMH NKKHTQGQFQKSSDFQKEELQQTFLPPERQG 158 1508 A 1939 1 423 TTHRLNVTAEPPCTSMPIYWMPDVPHRCTTA NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED RLHCQTQACPPLSWPQRLDILLGTARAIQFLH QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA RFSRFAGSSPIQSSM 159 1509 A 1974 3 401 HTSTARLLLHRGAGKEAVTSDGYTALHLAAR NGHLATVKLLVEEKADVLARGPLNQTALHL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR 160 1510 A 1982 2 417 KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA TYSEYCNNHPGACLELANLMKQGKYRHFFEA CRILQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRISHTGVKSTPGQSAAILMKLR	155	1505	A -	1920	-	360	
LLLKDMDLLRGREAELKQRVEAFESYQLELK DDYIIRTYRLIEDDRINIQISGHWQESP	100	1000	^	1747	-	203	
DDYIIRTYRLIEDDRINIQISGHWQESP	1 1						
1506							
DEDMHQKIAREMNLSETAFIRKLHPTDNFAQ RSCFGLIWFTPTTDLQILTSSILPSIL 157	156	1506	· <u>A</u>	1025	- - -	270	
RSCFGLIWFTPTTDLQILTSSILPSIL 157	150	1200	Λ.	1733	*	470	
157 1507 A 1936 584 305 ESKVNNEKFRTKSPKPAESPQSATKQLDQPTA AYEYYDAGNHWCKDCNTICGTMFDFFTHMH NKKHTQGQFQKSSDFQKEELQQTFLPPERQG 158 1508 A 1939 1 423 TTHRLNVTAEPPCTSMPIYWMPDVPHRCTTA NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED RLHCQTQACPPLSWPQRLDILLGTARAIQFLH QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA RFSRFAGSSPIQSSM 1599 A 1974 3 401 HTSTARLLHRGAGKEAVTSDGYTALHLAAR NGHLATVKLLVEEKADVLARGPLNQTALHL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDLAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR							
AYEYYDAGNHWCKDCNTICGTMFDFFTHMH NKKHTQGQFQKSSDFQKEELQQTFLPPERQG 158	157	1507	Δ	1026	584	305	l
NKKHTQGQFQKSSDFQKEELQQTFLPPERQG	157	1307	Λ	1730	J04	202	
158 1508 A 1939 1 423 TTHRLNVTAEPPCTSMPIYWMPDVPHRCTTA NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED RLHCQTQACPPLSWPQRLDILLGTARAIQFLH QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA RFSRFAGSSPIQSSM 1509 A 1974 3 401 HTSTARLLHRGAGKEAVTSDGYTALHLAAR NGHLATVKLLVEEKADVLARGPLNQTALHL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR							
NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED RLHCQTQACPPLSWPQRLDILLGTARAIQFLH QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA RFSRFAGSSPIQSSM 159 1509 A 1974 3 401 HTSTARLLHRGAGKEAVTSDGYTALHLAAR NGHLATVKLLVEEKADVLARGPLNQTALHL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR 160 1510 A 1982 2 417 KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDLAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR	159	1508	Δ	1030		423	
RLHCQTQACPPLSWPQRLDILLGTARAIQFLH QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA RFSRFAGSSPIQSSM 159 1509 A 1974 3 401 HTSTARLLHRGAGKEAVTSDGYTALHLAAR NGHLATVKLLVEEKADVLARGPLNQTALHL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR 160 1510 A 1982 2 417 KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR	170	1209	Λ	לכלו	1	423	
QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA RFSRFAGSSPIQSSM 159 1509 A 1974 3 401 HTSTARLLHRGAGKEAVTSDGYTALHLAAR NGHLATVKLLVEEKADVLARGPLNQTALHL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR 160 1510 A 1982 2 417 KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR							
RFSRFAGSSPIQSSM 159 1509 A 1974 3 401 HTSTARLLHRGAGKEAVTSDGYTALHLAAR NGHLATVKLLVEEKADVLARGPLNQTALHL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR 160 1510 A 1982 2 417 KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQQRHSHTGVKSTPGQSAAILMKLR							
159 1509 A 1974 3 401 HTSTARLLHRGAGKEAVTSDGYTALHLAAR NGHLATVKLLVEEKADVLARGPLNQTALHL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR 160 1510 A 1982 2 417 KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR							
NGHLATVKLLVEEKADVLARGPLNQTALHIL AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR 160 1510 A 1982 2 417 KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR	150	1500		1074	2	401	
AAAHGHSEVVEELVSADVIDLFDEQGLSALH LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR 160 1510 A 1982 2 417 KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR	לנו	1309	A	17/4	3	401	
LAAQGRHAQTVETLLRHGAHINLQSLKFQGG HGPAATLLR 160 1510 A 1982 2 417 KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR							
HGPAATLLR 160 1510 A 1982 2 417 KFLKDLEKQYNKEEPHILSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR							
160 1510 A 1982 2 417 KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR	1						
IYSEYCNNHPGACLELANLMKQGKYRHFFEA CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR	1.0	7.510		1000		1.00	
CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR	160	1510	A	1982	Z	417	
LKYTTQEHGDYSNIKAAYEAMKNVACLINER KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR] '						
KRKLESIDKIA 161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR]						
161 1511 A 1984 4 770 RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE LWLETLQGQRHSHTGVKSTPGQSAAILMKLR							
LWLETLQGQRHSHTGVKSTPGQSAAILMKLR	اا						
LWLETLQGQRHSHTGVKSTPGQSAAILMKLR SSHNASKTLNANNMETLECQSEGDIKEHPLL	161	1511	A	1984	4	770	
SSHNASKTLNANNMETLECQSEGDIKEHPLL							LWLETLQGQRHSHTGVKSTPGQSAAILMKLR
					<u> </u>		SSHNASKTLNANNMETLIECQSEGDIKEHPLL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq- uence	l	USSN 09/496	location	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	dence	•	914	correspondi ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline,
dena			714	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
ļ		İ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	1	! ,	peptide	sequence	/=possible nucleotide deletion, \=possible
	Ì			sequence		nucleotide insertion
	 		<u> </u>	Soqueio		ASCESEDSICQLIEVKKRKKVLSWPFLMRRLS
	1			-		PASDFSGALETDLKASLFDQPLSIICGDSDTLP
l l						RPIQDILTILCLKGPSTEGIFRRAANEKARKEL
ĺ		1	[KEELNSGDAVDLERLPVHLLAVVFKDFLRSIP
}			·			RKLLSSDLFEEWMGALEMQDEEDRIEALK
162	1512	A	1986	864	501	LLNSGLFSAPDGSNLEMRLTRGGNMCSGRIEI
						KFQGRWGTVCDDNFNIDHASVICRQLECGSA
1	}	1	}			VSFSGSSNFGEGSGPIWFDDLICNGNESALWN
		Ì				CKHQGWGKHNCDHAEDAGVICSSKD
163	1513	Α	2001	419	187	AVDLSIDESSLTGETTPCSKVTAPQPAATNGD
	!	1	}			LASRSNIAFMGTLVRCGKAKGVVIGTGENSE
	·					FGDIINLSTFVVHS
164	1514	A	2012	284	597	SLLCLFPGTSTVVCKPIVIETQLYVIVAQLFGG
						SHIYKRDSFANKFIKIQAIEILKIRKPNDIETFKI
]			ENNWYFVVADSSKAGFTTIYKWERETGFYSH
						QSFTR
165	1515	A	2013	2	403	EDPEELGHFYDYPMALFSTFELFLTIIDGPANY
i						NVDLPFMYSITYAAFAIIATLLMLNLLIAMMG
						DTHWRVAHERDELWRAQIVATTVMLERKLP
						RCLWPRSGICGREYGLGDRWILRVEDRQDLN
						RQRIQRYA
166	1516	Α	2019	2	927	CCQREGLGLKAVVQILLSHGRNGLPGEPASS
						QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF
ļ .]					NKNMVTERLQNVMVLEQCFSDSSSLYRFLTY
j :						SYLLAFNVWLLLAPVTLCYDWQVGSIPLVETI
						WDMRNLATIFLAVVMALLSLHCLAAFKRLE
j						HKEVLVGLLFLVFPFIPASNLFFRVGFVVAER
						VLYMPSMGYCILFVHGLSKLCTWLNRCGATT LIVSTVLLLLLFSWKTVKQNEIWLSRESLFRS
						GVQTLPHNAKVHYNYANFLKDQGRNKEAIY
1						HYRTALNNNKAWDYLCWRFRKTLTDLP
167	1517	A	2025	696	71	AAASAASSLTVTLGRLASACSHSILRPSGPGA
10,		•	2023	. 050	<i>'</i>	ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW
Į .						PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY
Į .						GRLFAVVHFASRQWKVTSEDLILIGNELDLA
						CGERIRLEKVLLVGADNFTLLGKPLLGKDLV
						RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV
						TTPQTVLRINSIEIAPCLL
168	1518	A	2046	2	366	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ
						RLQGAARVFMPLQAQVKAKASKPLQMQIKA
						PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS
						KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR
169	1519	A	2049	1	945	QNLEDREVLNGVQTELLTSPRTKDTLSDMTR
						TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI
]						EDNSRSKREGLFHENECIVKINNVDLVDKTFA
			l			QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS
						VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA
						NLTGTDSPETDASASLQQNKSPRVPRLGGKPS
						SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF
						TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ
						SGDRILEVNGRDVTGRTQEELVAMLRSTKQG
170	1500		2050	262	<u> </u>	ETASLVIARQEGHFLPRELVMFRSQSH
170	1520	Α	2050	363	1	PVATHLTKILNSDEHAVVISSAKTLCETVKDF
						VAKVEKTYDKTLENAVVADAVASKCSVLNE
						KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV
171	1521	A	2055	139	675	ESSSEESLGESKEQLGDDVTKPSSQKA
*''	1241	^ .	2000	137	0/3	IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG
						LDQPGCPKTILGGFQYFKCDHQRSVYLYLSSL
	L	Ц			<u> </u>	TO STATE OF WITHOUT ALL LEND OF STATE OF WITHOUT ALL LEND OF STATE

SEQ ID Mod m							· · · · · · · · · · · · · · · · · · ·
mucleotide seq	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence	ł .		hod				D=Aspartic Acid, E=Glutamic Acid,
			j				F=Phenylalanine, G=Glycine, H=Histidine,
172	eotide		1				
minn each residue of peptide residue of peptide residue of peptide sequence T-Threonine, V-Valine, W-Typtophan, Y-Tyrosine, X-Unknow, Y-Siop codon, Poposible nucleotide deletion, V-pos	seq-	uence			correspondi		M=Methionine, N=Asparagine, P=Proline,
	uence			914			Q=Glutamine, R=Arginine, S=Serine,
			İ				
				1		sequence	
RESCITTAYPOSYQDYRNGKCVSCOTSQKE		ł	ł	1		1	
SCPILGYYADNWKDHLRGKDPPMTKAFFDT				l	sequence	L	
		1		ł	}	ļ	RESCTITAYPCDSYQDYRNGKCVSCGTSQKE
172		1	ļ				
		1					
	172	1522	Α	2056	3	361	
HLVRIKEGRRRREFMARRILKCLKES 173 1523 A 2060 1 387 GTBLISMOJPFYGPPRTSENHAAAQVFALL QAYAFLQYLRDRLTKOEPOTLFFLGVSLAAA AVFLSVILLTYGYLAPWSGRFYSLWDTGVA KHIPIJASVSEHQPTTWVSFFFDLHILGCTFPA G							- ·
173		ļ	}	}	}		1 ·
174							
AVFLSYVILTYTGYIAPWSGRFYSLWDTGYA	173	1523	A	2060	1	387	
174		1	1	1			
1524			1		}]	
174		1	1	1	· ·		KIHIPIIASVSEHQPTTWVSFFFDLHILGCTFPA
RRILEBEEARLKYEKEEMERLEIORIBEKEKW HRILEAKDLERNREGE EIGHLEGCFFEAELK]]				
HRLEAKDLERRNEELEELYLLERCFFEAEKLK	174	1524	A	2071	74	443	LLMGPKAKKSGSKKKKVTKAERLKLLQEEEE
HRLEAKDLERRNEELEELYLLERCFFEAEKLK		1	1				
175	1		1		1		HRLEAKDLERRNEELEELYLLERCFPEAEKLK
AESSTVGWLCALFRYTHYGVGATGHGVVCG RRVLGLPSPAPMPMSPEGESRKEREVQ RRVLGLPSPAPMPMSPEGESRKEREVQ RRVLGLPSPAPMPMSPEGESRKEREVQ RRVLGLPSPAPMPMSPEGESRKEREVQ RRVLGLPSPAPMPMSPEGESRKEREVQ RLQPPYLEPGHELPATILLAFLAAV			i	1	ł	l	QETKLLSQWKHYIQCDGSPDPSVAQEMNT
AESSTVGWLCALFRYTHYGVGATGHGVVCG RRVLGLPSPAPMPMSPEGESRKEREVQ RRVLGLPSPAPMPMSPEGESRKEREVQ RRVLGLPSPAPMPMSPEGESRKEREVQ RRVLGLPSPAPMPMSPEGESRKEREVQ RRVLGLPSPAPMPMSPEGESRKEREVQ RLQPPYLEPGHELPATILLAFLAAV	175	1525	A	2083	139	486	AALTWSQPQEFWPMEMQPIVTDMVTVHWV
RRVLCGLPLPSPAPMPINISLEGESRKEREVQ RLOFPYLEPGHELPATTILAFLAAV		1		}			AESSTVGWLCALFRVTHVGVGATGHGVVCG
RI.QFPYLEPGHELPATTLLAFLAAV			ŀ		1	1	
1526		1	ł	1 .		1	
FOKFI-NILGDITITLKGWTGYRGGLDTKNDTT GIHSVYLVQGHEIMFHVSTMLPYSKENKQQ VERKRHGNDIVTIVPQGEESSPAFKPSMIRS HFTHIFALVRYNQQNDNYRLKIFSEESYPLFG PPLPTPLYFTDHQEFGESSPAFKPSMIRS HFTHIFALVRYNQQNDNYRLKIFSEESYPLFG PPLPTPLYFTDHQEFGDFLLVKLINGEKATLET PCI	176	1526	A	2092	3	587	
GIRSVYTVYQGHEIMFHVSTMLPYSKENKQQ	170	1320	1	2072	-	***	
VERKRHIGNDIVTIVFQEGEESSPAFKPSMIRS			i	1			
HFTHIFALVRYNQQNDNYRLKIFSEESPYLFG PPLPTPPFTDHQFFRDFLLVKLINGEKATLET PCI	1	1		ſ			
PPLPTPPVFTDHQEFRDFLLVKLINGEKATLET PCI	İ			1			
PCI	1	1		i		1	
177	1			1	Į	.	
CDGAWLAWACWYFGNDFPSPASAACSALLG CSVSTACLCVPLCSGSPLAFFRRTAALQEGLR RAYSVPLTLAETVASLWPALQELARCGNLAC RSDLQ 178 1528 A 2104 2 409 ALQSTLGAVWLGLLLNSLWKVAESKDQVFQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRILVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD 179 1529 A 2111 1 312 PITRSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLEGGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRNPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	177	1527	 _ _ _ 	2103	144	427	
CSVSTACLCVPLCSGSPLAPFRRTAALQEGLR RAVSVPLTLAETVASLWPALQELARCGNLAC RSDLQ 178 1528 A 2104 2 409 ALQSTLGAVWLGLLLNSLWKVAESKDQVFQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEPNTDKLIFGTGT RLQVFPNIQNPD 179 1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFJQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRNPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYLNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMOFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1''	1321	1	1 -103	1 ''	, ·'	
RAVSVPLTLAETVASLWPALQELARCGNLAC		i	1	1	Į.		
RSDLQ	1	ł	l		ł		
178		ł		1			-
PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRILVKGSKPSQQGRYNMTYERFSSSL LILQVFPNIQNPD 179 1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVTTQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICT 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACIST DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	179	1528	 	2104	12	400	
PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVFEADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD 179 1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNIAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	176	1520	^	2104	1	100	
LIILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD 179 1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		}	1	1	[
RLQVFPNIQNPD			1				
1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180	1	1	1	1	}		1 . ' - 2
MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	170	1520	 	2111	 , 	212	
SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQITEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1/9	1329	A	2111	1 '	312	
PPPLPLPACLG	1		1	1	1		
180		1		1	1		
LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	100	1500	1	13	12	266	
KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV	180	1530	A	2116	13	300	
QAFASVVCTFHLTACVSLHRIHNSTVV		1	1		1		
181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		1	1.	1	[
NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF			 	1	 	 	
DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	181	1531	A	2117	²	386	
GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1	1	1	1		1	
IS2 IS32 A 2123 I 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS IS33 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF			1		i		
182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		1	1		}	1	
GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF							
DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	182	1532	A	2123] 1	493	
LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		1	j				
IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF			1		1		
183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF			1	1			
183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1	+	1			1	
RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1		1				
RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	183	1533	A	2140	3	561	RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ
VYGIDKFMEDLKDMLGFAPSRYYYYMWKYI					1		RSGNYFVTMFDDYSATLPLLIVVILENIAVCF
	1	1	1	1	1	1	VYGIDKFMEDLKDMLGFAPSRYYYYMWKYI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
40		ł		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ	Ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
İ				peptide	- •	/=possible nucleotide deletion, \=possible
		1	ĺ	sequence		nucleotide insertion
						SPLMLLSLLIASVVNMGLSPPGYNAWIEDKAS
		1	l	ļ		EEFLSYPTWGLAVCASLDVFAILPVPVAFIGR
_						RFSLIDDGAGPFCSAAYTTTGCRTPYL
184	1534	A	2145	3	538	HELTVAAADRGQPPQSSVVPVTVTVLDVND
			ļ			NPPVFTRASYRVTVPEDTPVGAELLHVEASD
						ADPGPHGLVRFTVSSGDPSGLFELDESSGTLR
		Ì	ł		}	LAHALDCETQARHQLVVQAADPAGAHFALA
			i	Ī		PVTIEVQDVNDHGPAFPLNLLSTSVAENQPPG
						TLVTTLHAIDGDAGAFGRLRYHL
185	1535	A	2151	2	671	LDKLLDRMENYNIFNEYILKQVAATYIKLGW
		ļ				PKNNFNGSLVQASYQHEELRREVIMLACSFG
		ł				NKHCHQQASTLISDWISSNRNRIPLNVRDIVY
						CTGVSLLDEDVWEFIWMKFHSTTAVSEKKIL
]	ļ	LEALTCSDDRNLLNRLLNLSLNSEVVLDQDAI
				i		DVIIHVARNPHGROLAWKFFRDKWKILNTRI
		ļ				RQKTLEFDFAEPLILAFPIILYTAIDNPPLVREH
186	1536	A	2153	2	400	GPMCDKHSAFAEKFHAGFIDYTVHPLWETWA
100	1550	^	2133	2	700	HLALPDAQDILYTLEDNRNWVDSMIPQSPSPP
		1				LDEQNRDWQGLLENLHVELTLDEEDSEGPEK
		}	1	ł		EGEGQTYFTSSKTLCGIVPQNTDSLGETGIHIC
				1		AHDKSP
187	1537	A	2158	227	442	FNCFRVASDSFLENSSLLIMILPLRNATOEFIIR
		i			1	PGAVAYTCNPSTLGGWGGWITRSGVRDQPG
	Ì	ł		ł		OHGGTPS
188	1538	A	2167	3	486	AHLGGAWLTQRSLGSWAAPGPARAAKEVVA
	ļ	ľ	Ì			CIPQNQKMNIWRMKTSKHLQLLSFVLGAVSP
		Ì			Ì	AVVVPYMMVLQENGYGVEEGIPTLLMAASS
	1					MDDILAITGFNTCLSIVFSSGCARSSGSRNSKS
	ĺ	İ		!		LRTPLGTICEGCDDSSIFSHLDHSSKWSSTYG
				<u></u>		HSGA
189	1539	A	2168	2	412	EFLSSNQITQLPNTTFRPMPNLRSVDLSYNKL
		[(ĺ	QALAPDLFHGLRKLTTLHMRANAIQFVPVRIF
		1		ł		QDCRSLKFLDIGYNQLKSLARNSFAGLFKLTE
		1				LHLEHNDLVKVNFAHFPRLISLHSLCLRRNKV
100	1540		2170	64	200	AIVVSSLDW
190	1540	A	2179	64	399	MRLNQNTLLLESFGXXRPYTSEHAPTYHQW
		1	ļ			MKADELLRWTTSEPLTLEHEYAMQRTWLED AYECTFIVLDAEKRHAQPGATEESCMVGDVN
		1		1		LFLTDLEDLTLGEIEVLIAEP
191	1541	A	2190	1	469	CLDRAAGIRHERNVIYINETHTRHRGWLARR
	1541	l . .	,	Ι'	***	LSYVLFIQERDVHKGMFATNVTENVLNSSRV
		ļ ·			İ	QEAIAEVAAELNPDGSAQQQSKAVNKVKKK
	1	l	1	l	[AKRILQEMVATVSPAMIRLTGWVLLKLFNSF
			1			FWNIQIHKGQLEMVKAATETNLPLLFLPVHR
	1	ļ		1		SH
192	1542	A	2197	26	157	PSKXGGIRLLLTGTQLYGRFGSAIAPLGDLDR
					1	DGYNGEGREEPY
193	1543	A	2236	2	383	EYFPNSIWRSLFSTMDLGDIGFYTYRILQALS
				1		YTHSKGIMHRDVKPLNILCNSPRNKVILADW
						GLAEFYHPMRKYSVHVATRYYKSPEILLDYE
]]	}		1]	YYDYSLDIWAVGVILLELLTLKLHVFEGGDN
				<u> </u>		EQ
194	1544	A	2241	105	409	RKGVGKMPTSEGRPGQERSDWVTSYKVMGS
	1	1		1	}	NDSHTWVTVKNGSGDMIFEGNSEKEIPVLNE
1		i	I	1	1	LPVPMGARYIRINPQSWFDNGSICMRMEILGC
1	4	h.	ł	I	ł .	
105	1546		2245	1	670	PLPDPNNY
195	1545	A	2245	1	672	

000 10	L OP O TO	1.86.	LCCO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nuci-	peptide]	in USSN	location		l=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	İ			corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			914	ng to first		
	ļ .			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
	ĺ	l	1	residue of	sequence	
•			1	peptide	Ì	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
					ł	KLCGASSGIIDLLPSPSAATNWTAGLLVDSSE
					i	MIFKFDGRQGAKIPDGIVPKNLTDQFTITMW
			ļ.			MKHGPSPGVRAEKETILCYSDKTEMNRHHY
	l	1	ł	ł	ł	ALYVHNCRLVFLLRKDFDQADTFRPAEFHW
			1			KLDQQALAKVDGQPGKSITRQLQEMPVTIQG
						ISLKPS
196	1546	Α	2256	1	396	FRGTPVSGLTNRDTLAVIRHFREPIRLKTVKP
			ļ	l		GKVINKDLRHYLSLQFQKGSIDHKLQQVIRD
			l		1	NLYLRTIPCTTRAPRDGEVPGVDYNFISVEQF
	j	1	1	ļ		KALEESGALLESGTYDGNFYGTPKPPAEPSPF
						OPDPV
197	1547	A	2259	43	594	QLAIEIGVRALLFGVFVFTEFLDPFQRVIQPEEI
17,	***			1		WLYKNPLGQSDNIPTRLMFAISFLTPLAVICV
	1		1		[VKIIRRTDKTEIKEAFLAVSLALALNGVCTNTI
	l	1	1	1	1	KLIVGRPRPDFFYRCFPDGVMNSEMHCTGDP
	1	-	1		1	DLVSEGRKSFPSIHSSFAFSGLGFTTFYLAGKL
	1	-				HCFTESGRGKSWRLCAAILPL
100	1548	A	2275	3	404	TCTTVVVIPRMLVDFLSESKTISLPECATQMFF
198	1546	A	22/3	3	404	FLGFASNNCFIMAAMSYDRYTAIHNPLQYHT
		1	i	j		LMTRKICLQMMMASWMVGFLFSLCIIVTVFN
		}	1	j		LSLCDLNTIQHYFCDISPVVSLACNYTFYHEM
			ì		1	AIFVLSA
	L	<u> </u>	1000		300	LTOMFFIHALSAIESTILLAMAFDRYVAICHPL
199	1549	A	2315	1	375	DILAMATORI VALCATA
		í		1		RHAAVLNNTVTAQIGIVAVVRGSLFFFPLPLLI
			1	ł		KRLAFCHSNVLSHSYCVHQDVMKLAYADTL
						PNVVYGLTAILLVMGXDRMFISLSYFLII
200	1550	A	2334	2	409	PRVRPQQRKMSFFFKTELGEKLVTKFLFETDF
	1			İ		SDDPMLPSPDQLKKKAPFTNKKLKAHQTPVD
•	ļ					ILKQKAHQLASMQVQAYNGGNANPRPANNE
	1		1	{		EEEDEEDEYDYDYESLSDDNILEDRPENKSCH
	l			<u> </u>	l	DQLQFEYKEEM
201	1551	A	2350	3	512	ISWEAQIAEIIQWVSDEKDARGYLQALASKM
	1		1		ľ	TEELEALRSSSLGSRTLDPLWKVRRSQKLDM
				1		SARLELQSALEAEIRAKQLVQEELRKVKDAN
			į.	•		LTLESKLKDSEAKNRELLEEMEILKKKMEEK
}	ì	}	ł	1		FRADTGKLMLCDSALFEYKYFSNECFYFLFD
	1			1	}	LIVTLEAPTEFQIQY
202	1552	A	2351	1	1003	PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL
						NRVLERLAGGATRDSAASDILLDDIVLTHSLF
	1			}	ļ	LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA
۱.	1		1			CLAMLLHFLDTYQGLLQEEEGAGHIIKDLYL
	1		i	1		LIMKDESLYOGLREDTLRLHQLVETVELKIPE
			1	1		ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS
l	1	1	1	}		DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV
1	1		1	1	1	TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ
	1			1		PTEDCVFTALGINSHLFACTRDSYEALVPLPE
	}		1	1	1	EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR
	1	1	1	1		CVHELEFVDYVFHGE
	·	 	1-025	+	402	
203	1553	A	2361	2	403	NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH
	1	1	1	1	1	GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD
		1	1	1	1	WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL
]	1		1	1	LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ
1			<u> </u>	1	L	LGRITGLDP
	1554	A	2390	280	476	SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ
204	1				1	AGSRLGAMRRCAREMDATPMPPAPSCPSERV
204	1.55	-	1		1	
204						Т
204	1555	A	2400	543	745	T AAVALRDISWQQPYPMDFYAGSSLGPWTVN
		A	2400	543	745	Т

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
206	1556	A	2406	122	485	DLSPDSREDHPQGHRRLLPKRPVRGSLMPGH THHPCPVSSTTNDTPDQIWVSVGSLRMGTGG MGANASTSPRCWDLSSGNKKWIIQVPILASIV ESRGGLLATGVGGMCACVPRNQPLTGT
207	1557	A	2409	289	418	LWILYRHKQQVQHNHSNRLSCRPSQEDRAT HTIMVLDKENTLS
208	1558	A	2413	64	492	VQGTGXXFIAFTEAMTHFPASPVWAGMFFL MLINLGLGSMIGTMAGITTPIIDTFKVPKEMFT GGCCVFAFLVGLLFVQRSGNYFVTMFDDYSA TLPLTLIVILENIAVAWIYGTKKFMQELTEML GFRPYRFYFYMWKFVSP
209	1559	A	2417	3	877	EKERLLDEWFTLDEVPKGKLHLRLEWLTLMP NASNLDKVLTDIKADKDQANDGLSSALLILY LDSARNLPIRYKTNEPVWEENFTFFIHNPKRQ DLEVEVRDEQHQCPLGNLKVPLSQLLTSEDM TVSQRFQLGNSGPNSTIKMKIALRVLHLEKRE RPPDHQHSAQVKRPSVSKEGRKTSIKSHMSG SPGPGGSNTAPSTPVIGGSDKPGMEEKAQPPE AGPQGLHDLGRSSSSLLASPGHISVKEPTPSIA SDISLPIATQELRQRLRQLENGTTLGQSPLGQI QLTIP
210	1560	A	2422	35	456	REFAASDLEPFTPTDQPISPEAITQPSCIKRQRA AGNPGSLAATIDHKPCSAPLEPKIQASRNQRW GAVRAAESLTDIAEPASPQVHETPIDASQTQK VEPASKSRFTPELQAKVSHSRERALSTMDATP HHAQPQRGEG
211	1561	A	2431	1	764	RRYSOKLIOHTACOLLRTYPAATRIDSSNPNP LMFWLHGIOLVALNYQTDDLPLHLNAAMFE ANGGCGYVLKPPVLWDKNCPMYQKFSPLER DLDSMDPAVYSLTIVSGONVCPSNSMGSPCIE VDVLGMPLDSCHFRTKPIHRNTLNPMWNEQF LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL KALKRGYRHLOLRNLHNEVLEISSLFINSRRM EENSSGNTMSASSMFNTEERKCLQTHRVTVH GVPG
212	1562	A	2436	1	411	GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC TDLDHNLDKGYLTVLGEQATPTNRLGALPKG RANRTRDLELTYLAERIVRLTWIPGDANNRPI TDYDCQIEEHQ
213	1563	A	2445	1		MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT WLMPELHPKEQILELLVLEQFLSILPEELQIWV QQHNPESGEESVTLLEDLEREFDDPGQQVPAS PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG LPQEPSFRGISEHESNLVWKQGSATGEKLRSP SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKQHSSLT QHQRIHTGEKPYKCNQCGKAFSLRSYLIHQR IHSGEKAYECSECGKAFNQSSALIRHRKIHTG EKACKCNECGKAFSQSSYLIHQRIHTGEKPY ECNECGKTFSQSSKLIRHQRIHTGERPYECNE CGKAFRQSSELITHQRIHSGEKPYECSECGKA FSLSSNLIRHQRIHSG
214	1564	A	2461	1	615	GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL STRLVSVQEDAGKSPARNRSASITNLSLDRSG SPMVPSYETSVSPQANRTYVRTETTEDERKIL LDSVQLKDLWKKICHHSSGMEFQDHRYWLR THPNCIVGKELVNWLIRNGHIATRAQAIAIGQ

SEQ II NO: of	NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl- eotide	peptide seq-		in USSN	nucleotide location	location corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l delice	j	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	İ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1			peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			i	1		AMVDGRWLDCVSHHDQLFRDEYALYRPLQV
1	1.555		0464		2022	LFSVYCQLECSKLIL
215	1565	Α	2464	3	2932	GPGVRSSQDGMADVFVHLRTAWPRCSFISGQ HGPGRHGRRVCSSQDSMADVFVHLRTAWPT
	1	ļ		ļ		CSLISGOHGPGESVSYEDDDIPAPASLLHVNA
	i					AAPALTNPTAPVLCTAPNNTAQKEKVPSGMR
	ļ			1		QRPAGVRISSRTPDLTCAVSTHSTVPGVRISSC
	- 1			į		TPDLTCAVSIHSTVPSVCISSCTPDLTCAVSTH
ļ.	ļ			ł	•	STVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
	Ì				ļ	TPDLTCAVSIHATVPGVRISSCTPDLTCAVSIH
		ĺ		[ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
	1			}		TPDLTCAVSIHSTVPGVRISSCTPDLTCAVSIH
		1)	ļ]	ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
1		[TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
Ì						ATVPGVRISSCTPDLTCAVSIHATVPGVRISSC TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
	1	Ì	ł	ŧ	ı	ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
1				•		TPDLTCAVSIHATVPGVRISSCTPDLTCAVSTH
1			1			STVPGVRISSRTPDLTCAVSIHATVPGVHISSC
Í	Ì	ĺ		İ		TPDLTCAVSTHSTVPGVRISSRTPDLTCAVSIH
	ŀ		l			STVPGVCISSRTPDLTCAVSIHSTVPSVHISSCT
j		ļ	}	}		PDLTCAVSIHSTVPGVRISSRTPDLTCAVSTHS
İ			1			TVPGVHISSCTTDLTCAVSIHATVPGVHISSCT
ł	İ		į	l		PDLTCAVSTHTTVPGVRISSRTPDLTCAVSIHS
ļ	Į.	}	1	ł	1	TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT PDLTCAVSTHLTVPGVRISSRTPDLTCAVSIHA
İ				ĺ		TVPGVHISSCTPDLTCAVSIHATVPGVRISSRT
		•	1			PDLTCAVSIHATVPGVHISSCTPDLTCAVSTHS
ĺ	İ		{	1		TVPGVRISSRTPDLTCAVSIHSTVPGVHISSCT
1						PDLTCAVSTHSTVPGVHISSCTPDLTCAVSTH
J	1]]		STVPGVHISSRTPDLTCAVSIHATVPSVHISSC
			İ			TPDLTCAVSIHSTVPGLLTSVSQTSTG
216	1566	A	2477	1	414	FRTKSYRKGSYRCIVSEWIAEQGNWQEIQEK
1	ļ	ì	ł	l		AVEVATVVIQPTVLRAAVPKNVSVAEGKELD
1			1	1		LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS
			Ì	1		RVLARLDRDFLVHSSPHVALSHVDARSYHLL
217	1567	<u> </u>	2480	2	460	VRDVSKENSGYYY CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY
217	1567	A	2400	-	700	MQVLVCQHECVRELATRPGRLSPIENFLPLHY
]			1	1	}	DYLOFAYYRVGEYVKALECAKAYLLCHPDD
			1	1	1	EDVLDNVDYYESLLDDSIDPASIEAREDLTMF
				1		VKRHKLESELIKSAAEGLGXSYTEPNYW
218	1568	A	2483	140	383	AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT
						SANLLQLVRSSGDIQEGDLVEVVLSASATFED
L						LQIRPHALTVHSYRAP
219	1569	A	2489	3	428	SSRLVLLAGAAALASGSQGDREPVYRDCVLQ
			1	1		CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD
Ι,			J	}]	DCKYECMWVTVGLYLQEGHKVPQFHGKWP
						FSRFLFFQEPASAVASFLNGLASLVMLCRYRT
220	1570	<u> </u>	2409	1,	1207	FVPASSPMYHTCVAFAWVS
220	1570	A	2498	1	1297	MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH
	1		1			HHFGREAEMRELIQTFKGHMENSSAWVVEIQ
						MLKCRVDNVNSQLQVLGDHLGNTNADIQMV
1	}	}	1	1	}	KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL
ļ			1	1	}	KEDLEKADALTFQTLNFLKSSLENTSIELHVL
i		1		1		SRGLENANSEIQMLNASLETANTQAQLANSS
		1		1	1	LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS
i	1	I	1	-		LEGANAEIQGLKENLQNTNALNSQTQAFIKSS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion FDNTSAEIQFLRGHLERAGDEIHVLKRDLKM VTAQTQKANGRLDQTDTQIQVFKSEMENVN
						TLNAQIQVLNGHMKNASREIQTLKQGMKNA SALTSQTQMLDSNLQKASAEIQRLRGDLENT KALTMEIQQEQSRLKTLHVVITSQEQLQRTQ
221	1571	A	2501	3	500	RVRLNNDGLSPLMMAAKTGKIGIFQHIIRREV TDEDTRHLSRKFKDWAYGPVYSSLYDLSSLD TCGEEASVLEILVYNSKIENRHEMLAVEPINE LLRDKWRKFGAVSFYINVVSYLCAMVIFTLT AYYQPLEGTPPYPYRTTVDYLRLAGEVITLFT GVLFFFTN
222	1572	A	2508	3	395	DAHCQRKLAMQEFMEINERLTELHTQKQKL ARHVRDKEEEVDLVMQKVESLRQELRRTER AKKELEVHTEALAAEASKDRKLREQSEHYSK QLENELEGLKQKQISYSPGVCSIEHQQEITKL KTDLEKKS
223	1573	A	2544	-	412	NDPAIISNFSAAVVHTIVNETLESMTSLEVTK MVDERTDYLTKSLKEKTPPFSHCDQAVLQCS EASSNKDMFADRLSKSIIKHSIDKSKSVIPNID KNAVYKESLPVSGEESQLTPEKSPKFPDSQNQ LTHCSLSAA
224	1574	A	2552	401	1	GASLCFISTAFTVLTFLIDSCRFSYPERPIIFLSM CYNIYSIAYIVRLTVGRERISCDFEEAAEPVLI QEGLKNTGCAIIFLLMYFFGMASSIWWVILTL TWFLAAGLKWGHEAIEMHSSYFHIAAWAIPA VK
225	1575	A	2563	724	1	MSARKERREKGEEEGEGEKDGDEDEKEEEKE GLGEEEEKEAGKKKKKQEEKEKEKGAVYSR VARICKNDMGGSQRVLEKHWTSFLKARLNC SVPGDSFFYFDVLQSITDIIQINGIPTVVGVFTT QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT KTRVRYRLTAISVDHSAGPYH
226	1576	A	2571	449	3	EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP TRTVLTTDDISSSPIEDRDGRRGVAGNFPIFKV AGAACDRGMSLEACEAVTRKANRRTYTMG VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS
227	1577	A	2575	3	1197	VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS EDHINRKYAFKAAHPNMRTYYFCTDTGKEM ELWMKAMLDAALVQTEPVKRVDKITSENAP TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE KKALEAEKYGFQKDGQDRPLTKINSVKLNSL PSEYESGSACPAQTVHYRPINLSSSENKIVNVS LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS HRAQIMARYPEGYRTLPRNSKTRPESICSVTP STHDKTLGPGAEEKRRSMRDDTMWQLYEW QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI QRGDVTIDRRHRAHHPKVK
228	1578	A	2583	3	330	LPFLGLGSVLPQGMVMASPEMNPTICSVFEA HIVLLFHATIFRRGFQVTVLVGNVRQTAVVE KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD PTMGIKPHLWWVAA
229	1579	A	2589	1	448	DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR ECVAPNICKCKPGYIGSNCQTALCDPDCKNH GKCIKPNICQCLPGHGGATCDEEHCNPPCQH

C770 TD	1 000 00	137.	T 070	Don't seed	D - 2' 3 3	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	1	in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	ſ	USSN	location	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	1	09/496	correspondi		Q=Glutamine, R=Arginine, S=Serine,
uence	ĺ		914	ng to first	acid residue	T=Threonine, V=Valine, W=Tryptophan,
1				amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	1	ì	residue of	sequence	Y=1 yrosine, X=Unknown, Y=5top codon,
		l	}	peptide		/=possible nucleotide deletion, \=possible
		<u> </u>	ļ	sequence		nucleotide insertion
		1		1	l	GGTCLAGNLCTCPYGFVGPRCETMVCNRHC
	L	ļ.,		<u> </u>	100	ENGGQCLTPDICQCKPGWYGPTCSTA
230	1580	A	2593	2	138	AVTFSVVFAYVADITQEHERSMAYGLVCMFI
		ļ <u>.</u>			<u> </u>	LYLLYLLRNAFFLR
231	1581	A	2595	185	2	SGPYTDFTPWPTEEQKLLEQALKTYPVNPPER
						WEKIAEAVPGRTKKACIKRYKVADLRISK
232	1582	A	2596	1	391	STVTGQPRRLLDTAGHQQPFLELKIRANEPGA
	l	1			1	GRARRTPTCEPATPLCCRRDHYVNFQELGW
	1	1		1		RDWILLPEGYQLNYCSGQCPTHLAGSPGIAAS
		[1	1		FHSAVFSLLKANNPWPGRTSWCVPTARRPLS
			<u> </u>	L		LLYL
233	1583	A	2601	184	403	LLFSDEIIMAAPLRIADVTSGLIGGEDGRVYV
1	Į					YNGKETTLGDMTGKCKSWITPCPEEKVNVLQ
1	1		1			NSIPYWERIT
234	1584	A	2614	178	335	PLTLCLPENNKPPQADAVPDKELTLPVDSTTL
	1	1				DGSKSSDDQKIISYLWEKTQ
235	1585	A	2616	2	896	DVLEVYGTGVASTRHEMGTLDKHKELEDLV
	1					AKFLNVEAAMVFGMGFATNSMNIPALVGKG
1		ł				CLILRDEVNHTSLVLGARLLGATIGIFKHNYA
	1	· ·	1	ĺ		QSLEKLLRDAVIYGQPRTRRAWKKILILVEGV
1	1	1				YSMEGSIVHLPQIIALKKKYKAYLYIDEAHSI
1	i	1				GAVGPTGRGVTEFFGLDPHEVDVLMGTFTKS
		1	}	ļ		FGASGGYIAGRKARILSPPACLVPNTGSHSLH
Ì	1	1	1			RLTRDLOMNEAMVALVTDRLOGWNSGEGN
	1	i		į		WDRADKFGDLVDYLRVHSHSAVYASSMSPPI
ļ]	l				AEQIIRSLKLIMGLDGTTQ
236	1586	A	2621	1	392	NTSSFPAQPSSPARPSLPHLSQHPSNPLLPLAS
-50	1.500			1		ADHPQCGRFLPLHEPEPLCPSPSLSYPTLVSS
	J	j		į	}	WSSPFSSHHGCPPGLYPFPTSPKTIQPPGLAQL
	i	l			ļ	KMLCIPPGRQQLRGAQSMPGHGALSPLLLPP
				1	İ	A
237	1587	A	2628	398	1	DLVCKISGFGRGPRDRSEAVYTTMSGRSPAL
1237	1307	1 ^	2020	-/*	•	
1		1				
	ĺ	l	1	1		WAAPETLQFGHFSSASDVWSFGIIMWEVMAF
1						WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN
1						WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV
220	1500		2621	1	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV
238	1588	A	2631	1	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV
238	1588	A	2631	1	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ
238	1588	A	2631	1	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE
238	1588	A	2631	1	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW
238	1588	A	2631	1	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS
238	1588	A	2631	1	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI
238	1588	A	2631	I	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA
238	1588	A	2631	1	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT
238	1588	Ā	2631	1	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF
238	1588	. A	2631	1	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP
238	1588	A	2631	1	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS
						WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN
238	1588	A	2631	1	1104	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN
						WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW
						WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET
						WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ
						WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK
						WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK
						WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIINRE
239	1589	A	2636	1	678	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIFGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIIINRE KVNRDCI
						WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIIINRE KVNRDCI ELLDPTTPMRTKCIELLYAALTSSSTDQPKAD
239	1589	A	2636	1	678	WAAPETLQFGHFSSASDVWSFGIIMWEVMAF GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN LMHRLMLDCWQKDPGERPRFSQIHSILSKMV QDPEPPNV WSPCSLTCGVGLQTRDVFCSHLLSREMNETV ILADELCRQPKPSTVQACNRFNCPPAWYPAQ WQPCSRTCGGGVQKREVLCKQRMADGSFLE LPETFCSASKPACQQACKKDDCPSEWLLSDW TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA VVLRCPARRVRKPLITWEKDGQHLISSTHVT VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP KEALQTHKHQNGIFSNGSKAEKRGLAANPGS RYDDLVSRLLEQGAPCSSSKKKN MKPDNILLDEHGHVHITDFNIAAMLPRETQIT TMAGTKPYMAPEMFSSRKGAGYSFAVDWW SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ LSDVQNFPYMNDINWDAVFQKRLIFGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIIINRE KVNRDCI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EMANKELKQLRASYTESCIQEHYLPQVIDGTL Y
241	1291	A	2640	392	3	IRLTILRCVFMRLATICVLVFTLGSKITSCDDD TCDLCGYNQKLYPCWETQVGQEMYKLMIFD FIIILAVTLFVDFPRKLLVTYCSSCKLIQCWGQ QEFAIPDNVLGIVYGQTICWIGAFFSPLLPAM Y
242	1592	A	2642	405	1	YFKNTTLLLVGVICVAAAVEKWNLHKRIALR MVLMAGAKPGMLLLCFMCCTTLLSMWLSNT STTAMVMPIVEAVLQELVSAEDEQLVAGNSN TEEAEPISLDVKNSQPSVELIFVNEDILDFLMK SPLMISQACI
243	1593	A	2646	412	2	CLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAF LLNGSIDGIRHMFTPKLEIMLEPKVWREAATQ VFFALGLGFGGVIAFSSYNKRDNNCHFDAVL VSFINFFTSVLATLVVFAVLGFKANVINEKCIT QNSETV
244	1594	A	2650	1		MTTTLIGLLKTARLLRLVRVARKLDRYSEYG AAVLMLLMCIFALIAHWLACIWYAIGNVERP YLTDKIGWLDSLGQQIGKRYNDSDSSSGPSIK DKYVTALYFTFSSLTSVGFGNVSPNTNSEKIF SICVMLIGSLMYASIFGNVSAIIQRLYSGTARY HMQMLRVKEFIRFHQIPNPLRQRLEEYFQHA WTYTNGIDMNMVTNGTCSSCTSDDGHFILVS NHHQGGLIYSWNDAASMQRPFNHIKSSLLGS TSDSNLNKYSTINKIPQLTLNFSEVKTEKKNSS PPSSDKTIIAPKVKDRTHNVTEKVTQVLSLGA DVLPEYKLQAPRINKFTILHYSPFKAVWDWLI LLLVIYTAIFTPYSAAFLLNDREEQKRRECGY SCSPLNVVDLIVDIMFIIDILINFRTTYVNQNEE VVSDPASV
245	1595	A	2656	385	2	NLTWWPLFRDVSFYIVDLIMLIIFFLDNVIMW WESLLLLTAYFCYVVFMKFNVQVEKWVKQ MINRNKVVKVTAPEAQAKPSAARDKDEPTLP AKPRLQRGGSSASLHNSLMRNSIFQNKIHTLD PHV
246	1596	A	2660	200	506	VLVLQMNYYQMLIIYYVLFFKVNEFLAFEGPI LLDMRIKHLIKTNQLSQATALAKLCSDHPEIG IKGSFKQTYLVCLCTSSPNGKLIEEVSMFSFIS NYFLS
247	1597	A	2678	3	267	DAWVKNDIIFNQTERKQKISENLKHLASVRV VQKNLVFVVGLSQRLADPEVSPLVFFVILIFF VSLSYLEIIFDPAQLCDSSEHIIS
248	1598	A	2687	1	404	DFTTLAAMMRTLFSLFGDVRSDVHRFSVTLF GAAIKSVKNPDKKSIENQVLDSLVPLLLYSQD ENDAVAEESRQVLTICAQFLKWKLPREVYSK DPWHIKPTEAGTICRFFEKKCKGKINILEQTL MYSKNPKL
249	1599	A	2692	1	440	FRRRRRRERDCAAQGARRHCRHLAECKLV SFPIGIYKVLRNVSGQIHLITLANNELKSLTSK FMTTFSQLRELHLEGNFLHRLPSEVSALQHLK AIDLSRNQFQDFPEQLTALPALETINLEENEIV DVPVEKLAAMPALRSINL
250	1600	A	2693	459	21	LLPGSLGVPILHSQPWDPSPQCPHRAPSTPRRL PPLGALSQALTFLSRAAKNHSQDPGKGTKPFP AAPAAPPPRSSLPAPLPMGLKDKGPQPAPPTIF NSPWHPATLPGALGPQLSQAAPSPIPPPCLMG ISSCPDLKLTKSSTP
251	1601	A	2694	2	404	FVFDLKLRVPGFAALLIHGASSVPGPETVRLR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	l	1	ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
		ļ	ļ	sequence		nucleotide insertion
						QKRKKKAPDHSSGRKEELVTTHTVDKLETKK
		J]	}	J	PVGRVLCGLSGELLHSLLLPRRKTEKRALGSH
ĺ	ĺ	1	1			RKAGFPEHPVAPEPLSNSCQISKEGREQVLSEI
						GAGDCL
252	1602	Α	2697	421	1	PQKSHSGAYQCFATRKAQTAQDFAIIALEDG
	I	1				TPRIVSSFSEKVVNPGEQFSLMCAAKGAPPPT
	<u> </u>	1	İ		· ·	VTWALDDEPIVRDGSHRTNQYTMSDGTTISH
		ļ		Ĭ		MNVTGPQIRDGGVYRCTARNLVGSAEYQARI
		l	l			NVRGPPSIRAMRNIT
253	1603	A	2698	65	401	ACCOWRRTLIPAKSTTVSCTISTPHHPFRGSYS
		1	}	ł	}	FDDHITDSEALSRSSHVFTSHPRMLKRQPAIEL
	1	l		,		PLGGEYSSDVPRPLSTQLSSSLLGYFSTLMTG
	j	j			}	AAFTNNIASSTIIL
254	1604	Α	2699	438	301	GQIHSQDDPPFIDQLGFGVAPGFQTFVACQEQ
		1				RVRGPWEAGPGVGY
255	1605	A	2700	1	842	LQNREDSSEGIRKKLVEAEELEEKHREAQVS
	1			-		AQHLEVHLKQKEQHYEEKIKVLDNQIKKDLA
			}		i	DKETLENMMORHEEEAHEKGKILSEOKAMIN
1	J	į.	1	-		AMDSKIRSLEQRIVELSEANKLAANSSLFTQR
Ì	Į.				1	NMKAQEEMISELRQQKFYLETQAGKLEAQN
		ļ		1	}	RKLEEQLEKISHQDHSDKNRLLELETRLREVS
[1			ļ	LEHEEQKLELKRQLTELQLSLQERESQLTALQ
ì					ĺ	AARAALESQLRQAKTELEETTAEAEEEIQALT
		1		[ì	VGLGSNIFRLLKASARMSVELALSILAHP
256	1606	A	2701	2	405	FVGGPGADPPVAVMWDPRAARMDLTAYAE
250	1000	1 ^	2/01	1 2	1 403	LLKESGNQVLKNGNFSLAIRKYDEAIQILLQL
	[1				YOWGVPPRDLAVLLCNKSNAFFSLGKWNEA
-	i	ł	l	1		FVAAKECLQWDPTYVKGYYRAGYSLLRLHQ
		l				PYEAARMFFEGLR
257	1607	A	2702	2	399	FVESASSRPPGCFSGDGRFWLVSEGSRRGWD
231	1007	^	2/02	2	377	FNPSFSFLDPRYSVGGDENIGTVTTLANILREF
	1	İ	ļ		}	
1	1	i			ļ	NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE
	1	İ		İ	}	DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI GGNDL
250	1600	 	2200		1007	
258	1608	Α	2709	1	1097	SVGARQGEARDRIRRFFPKGDLEVLQAQVERI
	i .	ľ	1	ł	1	MTRKELLTVYSSEDGSEEFETIVLKALVKACG
	1	ŀ				SSEASAYLDELRLAVAWNRVDIAQSELFRGDI
	i	į.		ļ		QWRSFHLEASLMDALLNDRPEFVRLLISHGLS
]	}	}	1	}]	LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH
1		1		1	ļ	SAGTKAPALKGGAAELRPPDVGHVLRMLLG
			1	1		KMCAPRYPSGGAWDPHPGQGFGESMYLLSD
1	[i	[[[KATSPLSLDAGLGQAPWSDLLLWALLLNRA
1				1		QMAMYFWEMGSNAVSSALGACLLLRVMAR
		1	1.	Í	(LEPDAEEAARRKDLAFKFEGMGVDLFGECYR
		1		1		SSEVRAARLLLRRCPLWGDATCLQLAMQAD
L	1	I	İ	1	L	ARAFFAQDGVQSLPTQKWWGDMARR
259	1609	A	2721	1	403	VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR
l	1	1	1		1	GGFSQREMVTGERSPSPEEEEEEEEEGFGERA
	1	Ī				SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA
1	ļ	ļ])	PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE
		1				GGLRVRLP
260	1610	A	2728	1	477	LLGGDLRYHLQQNVHFTEGTVKLYICELALA
	1	1		[-		LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN
	1	1				IATVVKGAERASSMAGTKPYMAPEVFQVYM
[[1	i	[DRGPGYSYPVDWWSLGITAYELLRGWRPYEI
	1	1		Ţ		HSVTPIDEILNMFKVERVHYSSTWCKGMVAL
1		1		Ì		LRK
261	1611	A	2730	3	547	LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR
	1	1			1 "	EIYLQGCFKPLVSISPNDSLFEAVYTLIKNRIH
L				<u> </u>	<u></u>	1 COOLINDON DULL I IDUITION

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nuclcotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ţ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ	}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ł			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1 1	ľ	1	1	peptide	1 .	/=possible nucleotide deletion, \=possible
!			ļ	sequence		nucleotide insertion
						RLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLP
						RPSFLYRTIQDLGIGTFRDLAVVLETAPILTAL
						DIFVDRRVSALAVVNECGTHPQDERLGLGW
						GLGEPGSEERLFPAAITSR
262	1612	A	2733	3	431	GPEFPGSAKLVFLDLSYNNLTQLGAGAFRSA
			}		,,,,	GRLVKLSLANNNLVGVHEDAFETLESLQVLE
						LNDNNLRSLSVAALAALPALRSLRLDGNPWL
1						CDCDFAHLFSWIQENASKLPKGLDEIQCSLPM
1			}			ESRRISLRACRRPASRV
263	1613	A	2736	2	343	PARISGVDPPVRKATKGGENCSFEDNKNWOF
] 203	1013	^	2730	2	343	LWGLNGNFNFFKEPWGGRNNHAKGFRTTW
1						
				1		ARSSSQNNRTFQNNRNFLRLQRDSQKKGQFA
264	1614	Α	2738	2	245	RLISPLVNLPQSPGGLEFQYQAT
204	1014	Λ	2/30	2	243	RAMLKCLREGQPPPSYNWTRLDGPLPSGVRV
)						DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSH
265	1615	<u> </u>	2752			DTVDVLDPPEDSGKQVDL
203	1015	A	2752	2	388	AAGDAPLRSLEQANRTRFPFFSDVKGDHRLV
[•	LAAVETTVLVLIFAVSLLGNVCALVLVARRR
						RRGATACLVLNLFCADLLFISAIPLVLAVRWT
i i						EAWLLGPVACHLLFYVMTLSGSVTILTLAAV
	1616		0555			SLER
266	1616	Α	2755	192	1	AFREVGGYWGLLCEHLYAIPSKTSEGNWTAK
						LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP
	1616			12.		V
267	1617	Α	2760	434	714	ASRLEKQNSTPESDYDNTPNDMEPDGMGYM
						HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF
	1440					LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF
268	1618	A	2762	1	405	IACTFCGQDEWSPERSTRCFRRSRFLAWGEP
						AVLLLLLLSLALGLVLAALGLFVHHRDSPL
1						VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP
						ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE
						LPLSWAE
269	1619	A	2772	3	243	TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHL
						LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN
						LSKNLSFSEFCFDVSY
270	1620	A	2789	1	486	ELQSQQACTHTKETEQLRSQLQTLKQQHQQA
		ł	ŀ	ı		VEQIAKAEETHSSLSQELQARLQTVTREKEEL
		1				LQLSIERGKVLQNKQAEICQLEEKLEIANEDR
	j		J	}		KHALERFEQEAVAVDSNLRVRELQRKVDGIQ
						KAYDELRLQSEAFKKHSLDLLSKERELNGKL
						RHLSP
271	1621	A	2795	1	568	KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE
		ſ		ĺ		FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE
				.	ļ	RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN
1			}	ŀ	i	FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
		1			. 1	VKGIQEKOVFSNTKDSKOEITONKSFFSSVKE
		1	ļ			SQRDDGKGALNIVEFLRKREELHQILSTVKOP
272	1622	A	2797	8	523	KCMQGKYAGAMESEPCVCTEADFDCDYGYE
		- 1				RHSNGQCLPAFWFNPSSLSKDCSLGQSYLNST
		Í	ľ	i	ľ	GYRKVVSNNCTDGVREQYTAKPQKCPGKAP
]			RGLRIVTADGKLTAEQGHNVTLMVQLEEGD
1		1	į	ŀ		VQRTLIQVDFGDGIAVSYVNLSSMEDGIXHV
						YQNXGIXRXTVQVDNSLGS
273	1623	$\overline{\mathbf{A}}$	2801	72	395	HPSRSNVGPRQLTVWNTSNLSHDNRRKYIFS
	.023	.	2001	, 4		DEECONO! CIBIHODIDI BEBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB
		1	1	ļ		DEEGQNQLGIRIHQDIPLPPRRRELPALRTTNG
1	· .		-	ſ		KADSLNVSRNSVMQELSELEKQIQVIRQELQL
274	1624	A	2805	168	320	AVSRKTELEEYH
	1024	^	2007	100	340	ILWLYFETGTWVYPVFAKLSLLGLAALFSLRE IFIARNGVVGETLTHCKRV
					1	

NO. of No. of collaboration collaboration No. of collaboration no. of collaboration no. of collaboration no. of collaboration no. of collaboration no. of collaboration no. of collaboration no. of collaboration no. of collaboration no. of collaboration no. of collaboration no. of collaboration no. of collaboration no. of collaboration no. of collaborat	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
mucleotide seq-	NO: of						
Seg- wence			100		, ,		F=Phenylalanine, G=Glycine, H=Histidine,
1625 1626 1627 1627 1627 1628						1	
uence 914 ng to first amino acid residue of peptide residue of peptide peptide residue of peptide sequence peptide sequence Control of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence Control of Sequence sequence sequence Control of Sequence sequence sequence Control of Sequence sequence sequence Control of Sequence sequence sequence Control of Sequence sequence sequence Control of Sequence sequence sequence Control of Sequence sequence sequence Control of Sequence sequence sequence Control of Sequence sequence sequence Control of Sequence sequence sequence Control of Sequence sequence Control of Sequence sequence Control of Sequence sequence Control of Sequence sequence Control of Sequence sequence Control of Sequence sequence Control of Sequence sequence Control of Sequence sequence Control of Sequence sequence Control of Sequence sequence Control of Sequence sequence Control of Sequence sequence Control of Sequence sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence Control of Sequence		1 -		i	correspondi		M=Methionine, N=Asparagine, P=Proline,
mimo acid residue of peptide residue of peptide residue of peptide sequence	•		1	914		acid residue	Q=Glutamine, R=Arginine, S=Serine,
peptide		į				of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1625			1		peptide		/=possible nucleotide deletion, \-possible
MOKIFFQ			l	· _	sequence		
1626	275	1625	A	2812	208	321	,
			<u> </u>			L	
	276	1626	A	2813	41	266	
1627				1			
LFISYLHTPKIKQHEVLQAMGSLIGTIGEMED LFISYLHTPKIKQHEVLQAMGSLIGTIGEMED LFISYLHTPKITGKIKQ			L	-	ļ		
PLFQEEHGTATR WMTGWLEGGSXSVPKTPL	277	1627	A	2817	3	410	
GINOQPALNGSTSELPVKFLKTESLSSTLPTX		Į	1				
LPPHNSPGKIK		Į	ĺ	1		1	
1628		1	1		1	}	
VKLRLILHIEELQMEHDIRHYDLESVFMTWD PVDONPRLV	050	1600	ļ., —	2001	220	467	
PVDONPRLV	278	1528	A	2821	238	457	
1629			-				
TPPTFSSVPPPLPSLSSILHHSPLHSELNPHLQS	270	1620	<u> </u>	2022	242	 	
CRLPSRPSVSRELPPQSGPASSVPLAPTPLPDS	219	1029	A	2022	342	1,	
280		İ	1				
280			}	1			
CGQYWPLEKDSRIRFGFLTVTNLTGAVGEPG	280	1630	<u> </u>	2825	307	77	
281	200	1050	1 1	2023) 307	<i>} ''</i>	
1631			i			Ì	1
282	281	1631	A	2827	81	381	
NTTNMDEVPRPQALSGSSVVWVSGCVASRS VILSLTSG	201	1001	**	2021	"	301	
VILSLTSG		-		1	,		
TSSGKYNELGYPFGYLKASTTLTCVNLFVMP			1				
283	282	1632	A	2830	471	160	KLPXDKYELEPSPLTQYILERKSPHTCWQVFV
YLKTLPPYYL					ł		
1633		1	l				YNYPVLLPLLDDLFKVHKLKPNLKWRQAFDS
MKHYLHSQACSVFNYHLSPRTFPRYPGLMVP PLQCQMHPESTQFSIKLQPPPVGRKNRERVE SSEESAP			}	1			
PLQCQMHPEESTQFSIKLQPPPVGRKNRERVE SSEESAP	283	1633	A	2835	462	148	
SSEESAP			ł				
1634			'				
DYGLVVRGCLDLRYLAMRQRNNLLCNGLSL KSLAETVLNFPLDKSLLLRCSNWDAETLTED QVIYAARDAQISVALFLHLLGYPFSRNSPGEK KR				L			
KSLAETVLNFPLDKSLLLRCSNWDAETLTED	284	1634	A	2836	2	384	
QVIYAARDAQISVALFLHLLGYPFSRNSPGEK KR 285 1635 A 2843 20 271 PIRPYYSYSGLDRDCSWLPLAKAWLPDVMIL VCDRVSEDGINRQQAQEWCIKHGFELVELSP EELPEEDGKCLCVRKYGTYI 286 1636 A 2845 197 278 TAEDVLTVAYEHGYNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFTNSELHRAM NLHVGNLRLLSGPLDQVRAALPTPALSPKDK AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEQQPTPRDKLSCWYYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF				1			
KR			1		ĺ		
285			1	ł	1		1
VCDRVSEDGINRQQAQEWCIKHGFELVELSP EELPEEDGKCLCVRKYGTYI 286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFTNSELHRAM NLHVGNLRLLSGPLDQVRAALPTPALSPKDK AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEQQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	005	1625	 	0012		071	
EELPEEDGKCLCVRKYGTYI	285	1035	A	2843	20	2/1	
286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFTNSELHRAM NLHVGNLRLLSGPLDQVRAALPTPALSPKDK AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE EPGGELAETALHLAVRTADQTSLHLVE DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWYYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF			1	1			
287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFTNSELHRAM NLHVGNLRLLSGPLDQVRAALPTPALSPKDK AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE EPGQELAETALHLAVRTADQTSLHLVE DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	286	1635	-	2845	107	278	
NLHVGNLRLLSGPLDQVRAALPTPALSPKDK AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE EPGQELAETALHLAVRTADQTSLHLVE DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF							
AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE EPGQELAETALHLAVRTADQTSLHLVE DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	201	1037	^	2031	-	741	
QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF		}	1	1	}	}	
288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF			1	1			
288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF		1	-		l		
LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	288	1638	 	2850	12	469	
TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	200	1000	^	2007	-		
KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF					[
289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF		j	j]'	j		
289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF		1]	Ì	1		
DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	289	1639	A	2861	1 2	454	
DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	207	1	1]	1 -	1	
YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF			1	1	1		
290 1640 A 2868 1 378 FRQQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	ļ.				1		<u></u>
290 1640 A 2868 1 378 FRQQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF			1		1		
SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	290	1640	A	2868	1	378	
PDCASCLQAQDPLCGWCVLQGRCTRKGQCG	_		1		1		SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF
					1		PDCASCLQAQDPLCGWCVLOGRCTRKGQCG

000	CEO TO	1 1/	LCEC	David Co	l n-dia	
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	поа	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ł	USSN	location		
	uence		09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
seq- uence	uciice	[914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ucaicc	1		714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	l		}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	seductice	/=possible nucleotide deletion, \=possible
		l		sequence		nucleotide insertion
		├ ──	 	sequence	 	RAGQLNQWLWSYEEDSHCLHIQSLLPGHHPR
		l		Í	1	QE
291	1641	A	2870	<u> </u>	385	FRYMPNNRQQLLRKRHIGNDIVTIVFQEPGAL
271	1041	^	2070	'	363	PFTPKSIRSHFQHVFVIVKVHNPCTENVCYSV
			1	į		GVSRSKDVPPFGPPIPKGVTFPKSAVFRDFLL
		İ				AKVINAENAAHKSEKFRAMATRTRQEYLKD
						LA
292	1642	A	2877	3	188	RPTRPPPATTQSPESTMDTSLKKEKSAILDLYI
272	1042	^	2077	,	100	PPPPAVPYSPRYVAVHCHGMLVSCWCHL
293	1643	A	2878	 	427	REKEEEVEEEDKVVKETEKEAEOEKEEDSL
293	.1045	^	20/0	*	427	GAGTHPDAAIPSGERTCGSEGSRSVLDLVNYF
		ł	-	1	ł	LSPEKLTAENRYYCESCASLQDAEKVVELSQ
						GPCYLILTLLRFSFDLRTMRRRKILDDVSIPLL
		ļ		1		LRLPLAGGRGQAYDL
294	1644	A	2879	109	245	QLCCFCFRQTTLIVYILSFIGMVIFTFTLDLRYI
234	1044	^	2079	109	243	IIVFVTGGVLG
295	1645	A	2880	3	320	LASSQHGILNNLSLLFSICKTCIRTMDHHCPRA
293	1043	1 ^	2000	3	320	NNCVGEONHRFFCALHCKSKHFCIEFTLNTNF
-						FNCFLPGAEKSTIDAPFSLQPFLQDSKYNTALS
	ľ]	LSESISO
296	1646	Ā	2892	209	363	SQYSHSLDYHLLQVTKNPFTLGDSSNPGQTE
290	1040	^	2072	209	303	
297	1647	A	2893	8	424	RLQEFSQKMDQVRGHWPVST SPXTLXLDTFILLGIQDNILVLILATPPFMAGG
291	1047	Α	2073) °	424	KLYSTMGRFLRDRKNPACREMAVVLLANLA
		ł	1			
					•	QGDSLAARAIAVQKGSIGHLLGFLEDSLAAT QIQQSQASLLHMHNPPFEPTSVDMMRRACRA
		l		}	ļ	LLALAKVDDNHSEF
298	1648	A	2894	310	445	FWIYFPSFFMTGYLPLGFEFAVEITYPESEGTS
270	10.0	``	2074	310	1 342	SGLLNASAQVNL
299	1649	A	2898	<u> </u>	492	KIKAKNLTNYDLCSIFLGTSTLLVWVGVIRYL
277	1017	^	2070	*	752	GYFQAYNVLILTMQASLPKVLRFCACAGMIY
	ŀ					LGYTFCGWIVLGPYHDKFENLNTVAECLFSL
	ŀ	1		1	ł	VNGDDMFATFAQIQQKSILVWLFSRLYLYSFI
		-				SLFIYMILSLFIALITDSYDTIKKFQQNGFPETD
					ļ	LOEF
300	1650	A	2901	<u> </u>	445	PVWWNSLNGASEVTFSVHVKDGGSFPKTDST
				-		TVTVRFVNKADFPKVRAKEQTFMFPENQPVS
	[ſ		1		SLVTTITGSSLRGEPMSYYIASGNLGNTFOIDO
		l	1			LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP
	Ì	j	1			FSSYEKLDITVLDVNDNAPIF
301	1651	A	2902	162	433	THFICLPLGYCFPLLDKDLQLPSGFNCNFDFLE
					.55	EPCGWMYDHAKWLRTTWASSSSPNDRTFPG
		,				KPAVSEDMKELRPACSTYFNPRFPYKL
302	1652	A	2909	2	412	GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV
				-	***	EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI
		ł			ļ	CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN
		[-	1	SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ
						LVNRQDRAHFM
303	1653	A	2914	291	453	KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE
			1		.55	VPPTSILEHLQRRKIMKRPSSCS
304	1654	A	2926	179	354	PGVPSQALRKAESLKKCLSVMEAKVKAOTAP
	1057	``		*''	337	NKDVQREIADLGEVGAASLPPSSGPGA
305	1655	A	2938	135	438	GMGYLHAKGILHKDLKSKNVFYDNGKVVIT
505	1055	^	2,500	*33	000	DFGLFSISGVLQAGRREDKLRIQNGWLCHLA
		1		1		PEIROLSPDTEEDKLPFSKHSDVFALGTIWYE
[1	[1	[(LHAREWP
	l				1	LITTING VY I
306	1656		2044	2	320	
306	1656	A	2944	2	329	VRWNSCVNCSCAFGNGASLSTSLGESSGCLW
306	1656	A	2944	2	329	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	nou .	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
1	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence	i		714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	(i	ĺ	peptide	sequence	/=possible nucleotide deletion, \=possible
-						nucleotide insertion
}		 	ļ	sequence		SKTKLSFKVSSSA
307	1657	A	2950	2	411	NYLCIAKNSAGSAMGKTRLVVQVPPVIENGL
307	1037	^	2930	2	411	PDLSTTEGSHAFLPCKARGSPEPNITWDKDGQ
1	i			ĺ		PVSGAEGKFTIQPSGELLVKNLEGQDAGTYT
		1				CTAENAVGRARRRVHLTILVLPVFTTLPGDRS
	l	١.		i	ļ	LRLGDRLWLR
200	1/50	 	2951	1	407	PTRPPRVRFDNEFDAESQRKRTTSVSKMERM
308	1658	A	2931	1	407	DSSLPEEEEDEDKEAINGSGNAENRERHSESS
1			İ			
ļ	i			İ		DWMKTVPSYNQTNSSMDFRNYMMRDETLEP
		1	İ	1	ĺ	LPKNWEMAYTDTGMIYFIDHNTKTTTWLDP RLCKKAKAPEDC
	1.250		2054		170	
309	1659	A	2954	2	179	QDFLTLTLTEPTGLLYVGAREALFAFSMEALE
	<u> </u>					LQGAVRGGAVGGSRACQRARPRGAVLG
310	1660	Α	2959	1	419	QDMMERAIIDTFVGHDVVEPGSYVQMFPYPC
						YTRDDFLFVIEHMMPLCMVISWVYSVAMTIQ
				•		HIVAEKEHRLKEVMKTMGLNNAVHWVAWFI
]			}	TGFVQLSISVTALTAILKYGQVLMHSHVVIIW
L	<u> </u>					LFLAVYAVATIMFCF
311	1661	A	2963	3	465	MKPQMPGLGAPNGYGPGRGRAGVPGGPERR
1]	}			PWVPHLLPFSSPGYLGVMKAQKPGAGEGMK
		İ	İ		•	PQKPGLRGTLKPQKSGHGHENGPWPGPCNA
ŀ				1		RVAPMLLPRLPTPGVPSDKEGGWGLKSQPPS
	ļ					AVQNGKLPGHQPPNGYGPGAEPGFNGGLEPQ
	<u> </u>					KI
312	1662	A	2967	3	405	WLAQEWSPCTVTCGQGLRYRVVLCIDHRGM
ŀ	1	ļ				HTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV
1				İ		EAKLPWFKQAQELEEGAAVSEEPSFIPEAWS
	1			1		ACTVTCGVGTQVRIVRCQVLLSFSQSVADLPI
		<u> </u>			<u> </u>	DECEGPKPA
313	1663	A	2969	2	430	VVADNCRQGYLDALRFLERRGLTKEPVLWT
	į.					LVSKEPPAPADGNWDAGCDQRRKGGLSLNW
1						KVPHVQVKDVPNFEQLSPELEAALKKACTRD
1	ſ	(İ	PSRWARFWHSGPGQVLTYLLLPCTLPFEYTYF
		<u></u>		<u> </u>	<u></u>	RSRRLVVWLPDVPADLWWMQ
314	1664	A	2971	422	33	LDXSHNALQRLRPGWLAPLFQLRALHLDHNE
1				[1	LDALGRGVFVNASGLRLLDLSSNTLRALGRH
.]	1			Į	1	DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA
1		1		1	!	LSHLYLGCNELASFSFDHLHGLSATHLLTLDL
<u>L</u>		<u></u>	<u> </u>	<u> </u>		SSNRM
315	1665	A	2973	1	525	ITVSTHASGSPFGLEPQSGWLWVRAALDREA
		'		!		QELYILKVMAVSGSKAELGQQTGTATVRVSI
1		1		1	1	LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV
1	1			J	ļ	FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ
1		1		1		TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP
		<u>L</u>				RSTTGTVHVAVLDLNDNT
316	1666	A	2978	2	400	ELVVELVSAGKSGPERNTYEVQVVTGNVPKA
· ·						GTDANVYLTIYGEEYGDTGERPLKKSDKSNK
		1				FEQGQTDTFTTYAIDLGALTKIRIRHDNTGNR
		į.]]	}	AGWFLDRIDITDMNNEITYYFPCQRWLAVEE
L		L	<u></u> .	L		DDGQLSRE
317	1667	A	2981	3	440	VLNCQGRPTRPVRINGDGQEVLYLAESDNVR
	j]]	ļ	LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH
		}				HRENVFLSYQDKRINHGSLPHLQHRVRFAAS
1	1	1	1			DPSQYDASINLMNLQVSDTATYECRVKKTTM
])	· ·			ATRKVIVTVQARPAVPMCWTEGQ
318	1668	A	2995	119	414	LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA
1	1	1	1			ENYDARLLRIDIANTLREQVQELFNKTYGKQ
]		1		ŀ		RRTPGEGHVAAVDREVAGFPVPAEGISGETIH
319	1669	A	2999	2	332	GFFAYTYGRLVVVEDLHSGAQQHWSGHSAEI
L				·——		

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ł	l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
		<u> </u>		sequence		nucleotide insertion
						STLALSHSAQVLASASGRSSTTAHCQIRVWD
				İ		VSGGLCQHLIFPHSTTVLALAFSPDDRLLVTL
						GDHDGRTLALWGTGHL
320	1670	A	3000	693	322	IDESTGLIITVNYLDYETKTSYMMNVSATDQA
]		j]	PPFNQGFCSVYITLLNELDEAVQFSNASYEAA
			İ			ILENLALGTEIVRVQAYSIDNLNQITYRFDAY
201	1691	<u> </u>	2001			TSTQAKALFKIDAITVRGWGQGAPFFPI
321	1671	A	3001	6	383	RIPRGKACXTVLGRSTGELEGFASSRLPPQPC
	ŀ					GWGQSSDLLSRIDLDELMKKDEPPLDFPDTLE
l	ĺ	ĺ	1	1	ľ	GFEYAFNEKGQLRHIKTGEPFVFNYREHLHR
322	1672	A	3007	192	447	WNQKRYEALGEIITKYVYELLEKDCNSKKVS
322	10/2	A	3007	192	447	ERVRNSLFPGRGDSQCACCPSSPVWVFLETGF
						LFPWLFLQVEVIKKAYMQGEVEFEDGENGK
323	1673	A	3019	18	245	DGAASPRNVGHNIYILAHQLARH
323	10/3	^	3019	10	243	KELLFYHLIVNNINFFNTRYAKIHIPIIASVSEH
				[[QPTTWVSFFFDLHILVCTFPAGLWFCIKNIND
324	1674	A	3020	523	797	ERVFGKRGF
324	10/4	^`	3020	323	<i>'3'</i> ·	LCYFSARYHQRKIFGILYIFTLSAINRKEPNLFI YLFIFFEMESHSVTHAGVQRHNLNSLQPLPPG
		1				FKRFSCLCFLSSWNYRGAPPGPANF
325	1675	A	3022	2	156	NDFLPLYFGWVLTKKSSETLRKAGQVFLEEL
323	10/3	1.	3022	-	150	GNHKAFKKELRQCRWQVGAL
326	1676	A	3023	38	172	KMVRGSKKLISFFPGGPYGILAGRDPSKGLAT
320	****	^^	3023	1 30	172	FCLNKEALKDEFE
327	1677	A	3027	1	385	LTLEFLLLPAASELAHGKRLACCIVDHKLPEC
		- "		_	1	GFYGLYDKILLFKHDPTSANLLQLVRSSGDIQ
			,			EGDLVEVVLSASATFEDFQIRPHALTVHSYRA
			[PAFCDHCGEMLFGLVRQGLKCDGCGLNYHK
					·	RC
328	1678	Α	3030	13	569	ITRPTISCQRPGPGLAAGMLPYTVNFKVSART
						LTGALNAHNKAAVDWGWQGLIAYGCHSLV
		l	1			VVIDSITAQTLQVLEKHKADVVKVKWAREN
			1			YHHNIGSPYCLRLASADVNGKIIVWDVAAGV
				:		AQCEIQEHAKPIQDVQWLWNQDASRDLLLAI
						HPPNYIVLWNADTGTKLWKKSYADNILSFSF
						D
329	1679	A	3038	90	744	SVNLPPSLWPWEEAMDSTKSEPLKGSPEAED
8			1			GNIEYKKLVNPSQYRFEHLVTQMKWRLQEG
				· ·		RGEAVYQIGVEDNGLLVGLAEEEMRASLKTL
]]	J		HRMAEKVGADITVLREREVDYDSDMPRKITE
		l			-	VLVRKVPDNQQFLDLRVAVLGNVDSGKSTL
		İ				LGVLTQGELDNGRGRARLNLFRHLHEIQSGR
		1				TSSISFEILGFNSKGEVHGINGTQWGQTLRMG
220	1.000	 	2010		207	W
330	1680	A	3040	3	397	LCSTLLLLTIPSWVLSQITLKESGPTLMKPTET
		1				LTLTCTFSGFSLNTSGVGVAWIRQPPGKALE
		.				WLALIYWDDDKRYSPSLNDRLTIAKDTSRNQ
		1				VVLTMTNMGPVDTATYYCAQFARGARGSN
221	1/01	 	2042		1500	WFDPWGQ
331	1681	Α	3043	3	1509	AGIRHEAPPTTSNRHRRQIDRGVTHLNISGLK
		l				MPRGIAIDWVAGNVYWTDSGRDVIEVAQMK
						GENRKTLISGMIDEPHAIVVDPLRGTMYWSD
]			WGNHPKIETAAMDGTLRETLVQDNIQWPTG
		:				LAVDYHNERLYWADAKLSVIGSIRLNGTDPI
						VAADSKRGLSHPFSIDVFEDYIYGVTYINNRV
	-		7	o .		FKIHKFGHSPLVNLTGGLSHASDVVLYHQHK
					Ï	QPEVTNPCDRKKCEWLCLLSPSGPVCTCPNG
	1	1	1 :			KRLDNGTCVPVPSPTPPPDAPRPGTCNLQCFN
		ı				GGSCFLNARRQPKCRCQPRYTGDKCELDQC

No. of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence	NO: of	NO: of	hod	ID NO:	0 0		D=Aspartic Acid, E=Glutamic Acid,
1682 A 3045 3 3045 3 3 3 3 3 3 3 3 3			1				F=Phenylalanine, G=Glycine, H=Histidine,
1682	l .	_	1				1=ISOICUICINE, K-Lysine, L-Leucine,
amino acid or sequence residue of peptide residue of peptide sequence residue of peptide sequence residue of peptide sequence residue of peptide sequence residue of peptide sequence residue of peptide sequence residue of peptide sequence residue of peptide sequence residue of peptide resid	•	uence	ĺ				O=Glutamine R=Arginine S=Serine.
residue of peptide sequence	uence			914			T=Threonine V=Valine W=Tryptophan.
peptide	ĺ			1			
				i		0042000	/=possible nucleotide deletion, \=possible
WEHCRNGGICASSPSGMPTCRCPTGFTGPKC			1	ļ]	
FLÖBRCYYROSGYCENFÖTCÖMAADOSNO GROTATPEGSREPWIKGSRCLEGAGAVVNKQS GDVTCNCTDGRVAPSCLTCVGHCSNGGSCL MSKMMPECQCPEWIKGSRCLEGAGAVVNKQS GDVTCNCTDGRVAPSCLTCVGHCSNGGSCL MSKMMPECQCPEWIKGSRCLEGAGAVVNKQS GDVTCNCTDGRVAPSCLTCVGHCSNGGSCLVG MSKMMPECQCPPHMTGPRCEEHVFSQQQP GHLASILIP		-	 				WEHCRNGGTCAASPSGMPTCRCPTGFTGPKC
332		j]	j	ļ	<u>}</u>	FLGDRCQYRQCSGYCENFGTCQMAADGSRQ
MSISKMMPECQCPPHMIGTRCEEHVFSQQP GHLSsLIP						1	
332			l				
332			1]	
AVDEEYGDYFFEELDMLEESPFLKMTLPWGT			<u> </u>		<u> </u>		
LSSLRLQCRSQSDDGFIMWVRRGEQMIPTAD MPKSFFKRRRSMMEIKIN.QVFPVLRTSPREVLF EDRITAHADHVQQGFDWQSTAAVGVLKAV QFGEWSDQPRITKDVVCPHADFTDVVQRLQ LDLHEPPVSQCVQWVDBAKLNQMRREGIRY ARIQLCDNDIYFIPRNVHQFKTVVSAVCSLAW HIRLKQVHPVVEATQNTESNSNMDCGLTGKR ELEVDSQCVRIKTESEBACTEQLLTTASSSFP PASE	332	1682	A	3045	3	952	ATTISNEHTQVNRTYCCGTYRAGPMRQISLVG
MPKSFFKRRSMMEIKIN_QYI_PRISEPREVI_F	ļ		ļ				I SELDI OCDEOGDIOCDIMANADEGEOMIDIAD
BERTRAHADHUGGGFDWQSTAAVGVLKAV QFGEWSDQPRITKDVICFADPTIDVVQRLQ LDLHEPPVSQCVQWVDEAKLNQMRREGIRY ARIQLCDNDIYFIPRNVHQFKTVSAVCSLAW HIRLRQYHPVVEATQNITESINSNMDCGLTGKR ELEVDSQCVRIKTSESACTEIQLLTTASSSP PASE	•						MONGOER DED SYMPER VIL UNI DE L'ELEMENT L'AD
QFG@WSDQPRITKDVICFHAEDFTDVQRLQ LDLHEPPVSQCVQWVDDEAKINQMRREGIRY ARIQLCDNDIYFIPRNVIHQFKTVSAVCSLAW HRIKQYHFVVLATQNTESNISNINGCLTGKR ELEVDSQCVRIKTSEEACHEQLLTTASSSFP PASE ELYDSQCVRIKTSEEACHEQLTTASSSFP PASE SACSTGPELPGRATRSLTRPANQKGCDGDRL YYDGCAMIAMIGSVFAQGSOFSLDDVEVLT ATLDLEDVRSYRAEISSRNIAVSAPVDTCVG CSSKTWKVAFFVRAWWRP SACSTGPELPGRATRSLTRPANQKGCDGDRL YYDGCAMIAMIGSVFAQGSOFSLDDVEVLT ATLDLEDVRSYRAEISSRNIAVSAPVDTCVG CSSKTWKVAFFVRAWWRP SACSTGPELPGRATRSLTRPANQKGCDGDRL YYDGCAMIAMIGSVFAQGSOFSLDDVEVLT ATLDLEDVRSYRAEISSRNIAVSAPVDTCVG CSSKTWKVAFFVRAWWRP SACSTGPELPGRATRSLTRPANQKGCDGDRL SACSTGPELPGRATRSLTRPANQKGCDGBLCSKTWKVAFFVRAWWRP SACSTGPELPGRATRSLTRPANQKGCDGLGSKTWAFTVASAPVALRCH SACSTGPELPGRATRSLTRPANQKGCDGSTC CSSKTWKVAFFVRAWWRP SACSTGPELPGRATRSLTRPANQKTGPTGRATGALTCH ARTOLD CSSKTWKVAFFVRAWWRP SACSTGPELPGRATRSLTRPANQKTGPTGRATGALTCH CSSKTWKVAFFVRAWWRP SACSTGPELPGRATRSLTRPANQKTGPTGRATGALTCH CSSKTWKVAFFVRAWWRP SACSTGPELPGRATRSLTRPANQKTGTGRATGATT CSGGG SASSLRRCLTGR NCEGQIRTYKTCSNHDCPJAEDFTRAQCSA YNDVQYQGHYYEWLPR YNDPAAPCALKCH AQGNILVVELAFKVLOGGGSTC RLYRGQSKSHVSPERREENVIAVPLGSRSVRI TVKGPAHLFESKTLQGSKGEHSFNSPGVFVV ENTTVEPQRGSERQTFKIPGPLMADFIFKTRY TAKADSVVQFFFYQPISHQWRQTDFFFCTVT CGGG CSGGGGGGGGGGGGGGGGGGGGGGGGGGGG	l		ł				
LDLHEPPVŠQCVQWVDEAKLNOMREGIRY	i		1				
ARIOLCDNDÍYPIPRNYIHOFETYSAVCSLAW HIRLKOYHPVVEATOPENSNMDCGLTĞKR ELEVDSQCVRIKTESESNMDCGLTĞKR ELEVDSQCVRIKTESESNMDCGLĞLTĞKR ELEVDSQCVRIKTESEEACTEIQLLTTASSSFP PASE	1		1				
HIRLKQYHPVVEATQNTESNSNMDCGLTGKR ELEVDSQCVRIKTESEEACTEIQLLTTASSSFP PASE			1	1	}	l	
		Ì	j	1			HIRLKQYHPVVEATQNTESNSNMDCGLTGKR
333	1	· '		1			
YYDGCAMIAMNGSVFAQGSQFŠLDDVEVLT ATLDLEDVRSYRAGISSRNLAVSAPVDTCVG CSSKTWKVAFFVRAWRP			1				
ATLDLEDVRSYRAEISSRNLAVSAPVDTCVG	333	1683	Α	3046	497	167	
CSSKTWKVAPFVRAWWRP		1		ĺ		1	YYDGCAMIAMNGSVFAQGSQFSLDDVEVLT
334			1				
EASGANDEIVQLRSEVDHLRREITEREMQLTS QKQVRRVNKVVRSLEDF QKQVRRVNKVVRSLEDF QKQVRRVNKVVRSLEDF NCBQMRYKTCSNIEDCPPDAEDFRAQQCSA YNDVQYQGHYYEWLPRYMDPAAPCALKCH AQGMLVVELAFKVLDGTRCMTDSLDMCISG ICQAVGCDRQLGSNAKEDNCGVCAGDGSTC RLVRQGSKSHVSPEKREENVLAVPLGSRSVRI TVKGPAHLFIESKTLQGSKGEHSFNSPGVFVV ENTTYEFQRGSERQIFKLIPGPLMADFIFKTRY TAAKDSVVQFFFYQPISHQWRQTDFFPCTVT CGGG VVGKQEAGAHSDSCCLLHTPPRLTPAHSRKA LRNSRIVSQKDVHVCIMCLRAIMNYQVSRG AWDWRLGSPACPHWGLHKLPRLWDPLSLYP VLCWGT UTSILVELTRFETLTPRFSATVPPCWVEVQQE QQQRRHPQHLHQQHHGDAAQHTRTWKLQT DSNSWDEHVFELVLPKACMVGHVDFKFVLN SNINTNPQIQVTLLKNKAPGLGKVNGLRLCFF LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLHVKAVNERGTEE CNGGMPPVVLPSKKYNLSKAPGLGKVNGLRLCFF LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLHVKAVNERGTEE CNGGMPVVVLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK SAFYNYHVLELLQMLVTGGVSSQLEQHLDK DKYYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVMPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP SSVTTMLSWV PSSVTTMLSWV PSSVTTMLSWV PSSVTTMLSWV SSVTTMLSWV SV					<u> </u>		
QKQVRYNKVVRSLEDF	334	1684	A	3053	37	276	
335		1	1	İ			
NCEGQNIRYKTCSNHDCPPDAEDFRAQQCSA	1225	1605	 	2054	<u> </u>	016	WDAWGDWSDCSRTCGGGASVSI RRCI TGR
YNDVQYQGHYYEWLPRYNDPAAPCALKCH AQGQNLVVELAPKVLDGTRCNTDSLDMCISG ICQAVGCDRQLGSNAKEDNCGVCAGDGSTC RLVRGQSKSHVSPEKREENVIAVPLGSRSVRI TVKGPAHLFIESKTLQGSKGHSFNSPGVFVV ENTTVEFQRGSERQTFKIPGPLMADFIFKTRY TAAKDSVVQFFFYQPISHQWRQTDFFPCTVT CGGG	333	1083	A	3034	2	040	NCEGONIR YKTCSNHDCPPDAEDERAOOCSA
AQGQNLVVELAPKVLDGTRCNTDSLDMCISG ICQAVGCDRQLGSNAKEDNCGVCAGDGSTC RLVRGQSKSHVSPEKREENVIAVPLGSRSVRI TVKGPAHLFIESKTLQGSKGHSFNSPGVFVV ENTTVEFQRGSERQTFKIPGPLMADFIFKTRY TAAKDSVVQFFFYQPISHQWRQTDFFPCTVT CGGG				1	1		YNDVOYOGHYYEWLPRYNDPAAPCALKCH
ICQAVGCDRQLGSNAKEDNCGVCAGDGSTC	1	1	İ				
TVKGPAHLFIESKTLQGSKGEHSFNSPGVFVV		1					
BNTTVEFQRGSERQTFKIPGPLMADFIFKTRY	1	1	1)	1	ļ	RLVRGQSKSHVSPEKREENVIAVPLGSRSVRI
TAAKDSVVQFFFYQPISHQWRQTDFFPCTVT CGGG 336 1686 A 3058 54 347 VVGKQEAGAHSDSCCLLHTPPRLTPAHSRKA LRNSRIVSQKDDVHVCIMCLRAIMNYQVSRG AWDWRLGSPACPHWGLHKLPRLWDPLSLYP VLCWGT 337 1687 A 3059 2 709 ILTSLVELTRFETLTPRFSATVPPCWVEVQQE QQQRRHPQHLHQQHHGDAAQHTRTWKLQT DSNSWDEHVFELVLPKACMVGHVDFKFVLN SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK 338 1688 A 3060 85 384 KAFYNYHVLELLQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP 339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 1249 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASSPYPKRKQICELLLRKGGNINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL		'			İ		
336					1		
336			1]			1
LRNSRIVSQKDDVHVCIMCLRAIMNYQVSRG AWDWRLGSPACPHWGLHKLPRLWDPLSLYP VLCWGT 337 1687 A 3059 2 709 ILTSLVELTRFETLTPRFSATVPPCWVEVQQE QQQRHPQHLHQQHHGDAAQHTRTWKLQT DSNSWDEHVFELVLPKACMVGHVDFKFVLN SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK KAFYNYHVLELQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP 339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 1249 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL							
AWDWRLGSPACPHWGLHKLPRLWDPLSLYP	336	1686	A	3058	54	347	
337 1687 A 3059 2 709 ILTSLVELTRFETLTPRFSATVPPCWVEVQQE QQQRRHPQHLHQQHHGDAAQHTRTWKLQT DSNSWDEHVFELVLPKACMVGHVDFKFVLN SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK 338 1688 A 3060 85 384 KAFYNYHVLELLQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP 339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 1249 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVKHEAKVNAL			1		1		
337 1687 A 3059 2 709 II.TSLVELTRFETLTPRFSATVPPCWVEVQQE QQRRHPQHLHQQHHGDAAQHTRTWKLQT DSNSWDEHVFELVLPKACMVGHVDFKFVLN SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK 338 1688 A 3060 85 384 KAFYNYHVLELLQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP 339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL		1	1	1	1		
QQQRRHPQHLHQQHHGDAAQHTRTWKLQT DSNSWDEHVFELVLPKACMVGHVDFKFVLN SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK VAAGKEKSSNVKNENTSGTRK KAFYNYHVLELLQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP 339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL	227	1697	 	3050	+,	709	
DSNSWDEHVFELVLPKACMVGHVDFKFVLN SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK 338 1688 A 3060 85 384 KAFYNYHVLELLQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP 339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL	337	100/	^	2027	1~	"	
SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK 338 1688 A 3060 85 384 KAFYNYHVLELQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP 339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL				1		1	
LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK 338 1688 A 3060 85 384 KAFYNYHVLELLQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP 339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 1249 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL	1	1	1	1	1		
CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK 338 1688 A 3060 85 384 KAFYNYHVLELLQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP 339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 1249 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL	}		İ	1			LEDHKEDILCGPVWLASGLDLSGHAGMLTLT
VAAGKEKSSNVKNENTSGTRK 338 1688 A 3060 85 384 KAFYNYHVLELLQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP				1		1	
338 1688 A 3060 85 384 KAFYNYHVLELLQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP				1	1	1	
DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP 339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 1249 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL					<u> </u>	<u> </u>	
SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP 339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 1249 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL	338	1688	Α	3060	85	384	
BELNP 339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 1249 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL				1			
339 1689 A 3063 236 362 CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV 340 1690 A 3065 3 1249 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL		}		1			·
PSSVTTMLSWV 340 1690 A 3065 3 1249 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL	222	1.000	1	2002	1226	1262	
340 1690 A 3065 3 1249 DLWQFTPLHEAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL	339	1689,	A	3063	236	362	
LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL	240	1.000	+	2065	+	12/0	
QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL	340	1090	I A	2002	,	1249	
LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL		1	1	1	}	1	
			1 .				
			1	ļ	1		
					1		

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IISLQGFTALQMGNENVQQLLQEGISLGNSEA DRQLLEAAKAGDVETVKKLCTVQSVNCRDIE GRQSTPLHFAAGYNRVSVVEYLLQHGADVH
						AKDKGGLVPLHNACSYGHYEVAELLVKHGA VVNVADLWKFTPLHEAAAKGKYEICKLLLQ HGADPTKKNRDGNTPLDLVKDGDTDIQDLLR GDAALLDAAKKGCLARVKKLSSPDNVNCRD TQGRHSTPLHLAGK
341	1691	A	3070	1	547	GVLIPSFQNQLFADILAGIESVTSEHNYQTLIA NYNYDRDSEEESVINLLSYNIDGIILSEKYHTI RTVKFLRSATIPVVELMDVQGERLDMEVGFD NRQAAFDMVCTMLEKRVRHKILYLGSKDDT RDEQRYQGYCDAMMLHNLSPLRMNPRAISSI HLRMQLMRDALSANPDLDGVFCTN
342	1692	A	3073	463	3	RINRCRKPSDADILVPGDTISLIGTTSLRIDYNE IDDNRVTAEEVDILLREGEKLAPVMAKTRILR AYSGVRPLVASDDDPSGRNVSRGIVLLDHAE RDGLDGFITITGGKLMTYRLMAEWATDAVC RKLGNTRPCTTADLALPGSQEPAKVP
343	1693	A	3075	250	1	LLIYLAIFAPVAMSALAGVKSVQQVRIRAAQS LGASRAQVLWFVILPGALPEILTGLRIGLGVG WSTLVAAELIAATRGLGFM
344	1694	A	3076	2	138	LYFDAYLQSLQVAAISTFCCLLIGYPLAWAV AHSKPSTRNILLLL
345	1695	A	3078	469	3	LKIRGQRIELGEIDRVMQALPDVEQAVTHAC VINQAAATGGDARQLVGYLVSQSGLPLDTSA LQAQLRETLPPHMVPVVLLQLPQLPLIANGKL DRKALPLPELKAQAPGRAPKAGSETIIAAAFS SLLGCDVQDADADFFALGGHSLLAMKLAT
346	1696	A	3082	404	2	QNITSKDLDVRLDPQTVPIELEQLVLSFNHMI ERIEDVFTRQSNFSADIAHEIRTPITNLITQTEI ALSQSRSQKELEDVLYSNLEELTRMAKMVSD MLFLAQADNNQLIPEKKMLNLAHEVGKVFD QFEALPE
347	1697	A	3084	3	340	NELTFKEAEISKLYTKVHPAYRTLLEKRQALE DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ QVYMQLLNKEQELKITEASTVGDVRIVDPAIT QPGVLKPKKGLIILGAI
348	1698	A	3086	723	10	TQAMVWQQKACAEDDPQLSGRHWLHAATL YNIAAYPHLKGDDLAEQAQALSNRAYEEAA QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI AMLTIDMPSVGFSSKWKLTQDSSLLHQHVLK ALPNVPWVDHTRVAAFGFRFGANVAVRLAY LESPRLKAVACLGPVVHTLLSGLKCQQQVPE MYLDVLASRLGMHDASTKSSTRENH
349	1699	A	3087	2	249	RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV GFGVAMSQALGPFSLRAGVASSTLGIAQVCG SSLWIWLAAVVGIGAWNM
350	1700	A	3099	3	424	EAPEATPQPSQPGPSSPISLSAEEENAEGEVSR ANTPDSDITEKTEDSSVPETPDNERKASISYFK NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN KDTVIIVSEPSEDEESQGLPTMARRNDDISELE DLSGMEDLK
351	1701	A	3108	2	404	IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG MFGIVVMVIETELSWGAYYKAPLYSLALKCL ISLFTIILLGLTIVYHAREIQLFMANYGADDWR SALTYEPIFLILLEALRGVIHATPCRVSLSLWD GLDLP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	_	/=possible nucleotide deletion, \=possible
ļ	ļ	ļ		sequence		nucleotide insertion
352	1702	A	3110	341	2	AQLAEVCPPQTLLTTNTSSISITALAAEIKNPER
						VAGLHFFNPAPVMKLVEVVSGLATAAEVVE
		<u> </u>				QLCELTLSWGKQPVRCHSTPGFIVNRVARPY
}	Ì		İ	l		YSEAWRALEEQVAAPEVI
353	1703	A	3111	3	188	HFSLFRIAFAVFLTYMTVGLPLPVIPLFVHHEL
]					GYGNTMVGIAVGIQFLATVLTRGYAGRLA
354	1704	A	3116	367	225	WQLFHLNGTFLNIGETDTESCVNGWVYDRSS
1	}	}	1	ł		FPFSNMTEVRGLVFLS
355	1705	A	3117	101	53	VINLVYLISSPRPELKPVDKESEVVMKFPDGF
1	•		}		1	EKFSPPILQLDEVDFYYDPKHVIFSRLSVSADL
j .	1		1	ļ	}	ESRICVVGENGAGKSTMLKLLLGDL\APVRGI
l		ļ			i	RHAHRNLKIGYFSQHHVGAAGT*TFSACGNL
	ł				}	LGTQVFLGRPEEEY\RHQLGFGMGISGELGHA
j	1]	j	}	}	SSLPACLGGQKEAEVAFCSDGLLPCPNFL\IL\
	1	İ	ļ	1		DEPTN\HLGHGRAIEALGPCLQTISGVGVILVS
}			1		ļ	HE*SALSRLVCRE\LWVC*GRSTSPF
356	1706	A	3121	137	466	RGGRDWGEHNQRLEEHQARAWQGAMDAG
						AASREHARWQGTGLAPGTRVAVAPTCVQGL
		l	ļ	}		PQERSVCRPFFSSRWREGPVWALGAGAHGKP
		1	į			RWSGGVRCVVRGGRWFTPAPH
357	1707	A	3124	1249	229	MLEAPGPSDGCELSNPSASRVSCAGQMLEVQ
ŀ		Į	ļ			PGLYFGGAAAVAEPDHLREAGITAVLTVDSE
1		Į				EPSFKAGPGVEDLWRLFVPALDKPETDLLSH
ľ	Ì	1	ľ	Ì		LDRCVAFIGQARAEGRAVLVHCHAGVSRSV
		1	}	ļ	j	AIITAFLMKTDQLPFEKAYEKLQILKPEAKMN
1		ŀ				EGFEWQLKLYQAMGYEVDTSSAIYKQYRLQ
ļ		1	1			KVTEKYPELQNLPQELFAVDPTTVSQGLKDE
		1	İ]	VLYKCRKCRRSLFRSSSILDHREGSGPIAFAH
1		1.				KRMTPSSMLTTGRQAQCTSYFIEPVQWMESA
ļ	1.	1	i		l	LLGVMDGQLLCPKCSAKLGSFNWYGEQCSC
		İ				GRWITPAFQIHKNRVDEMKILPVLGSQTGKI
358	1708	A	3127	816	139	EVETLGPRTPGP/EAQSPTPGSCPGWQEPSPGP
· ·		ļ	1		1	TPPP*LSGPGPQGAPVLGKLLPDPEETPAGKTP
				ļ		LGKHFWWGL\PVTSANFSPGAAA*FGGALSPP
1		l				GGDL/GHMLLQGPPSPFRLQQQ*QTPPGSHSP
)]		}	PTANREINPGPAAAADTRSCWGHKRSWRGW
				i		RGLAPWRLGFGSPGIP*PAPAGIP/GRPTWEGG
						KGAGGKPSETLTRSPPVWRGKRGSANGFLSW
	<u> </u>					VQILQ
359	1709	A	3132	3	191	HEHLLLLLCVFLVKSQGVNDNEEGFFSARG
,		1	1		1	HRPLDKKREDAPNLRPALAD\ITVCDYRAQIA
	<u> </u>	<u> </u>		L	<u> </u>	*AASTPKRAASIAHNAVSCR*AQIA
360	1710	A	3134	1	286	REPPRPALLFF*DRVSLCCPGWNAVVQSQLT
1		1	1	1	1	AAPTSQVQ/SDSPTFPSSWDYRHVPEYPANFL
	1					*RQGFPMLPRLVSNSWAQTVHPPRPPKVLDL
		<u> </u>				QA
361	1711	A	3135	56	1449	PVPAPRVSPSARGAPGRPRLPGVRGPRHS/WA
	1	İ		ł		AD*RGSRM/PPRAPAPSPTGP/APGGKKVRGR
1		1				VPEDPDAYEPRCSAL*V*PTHVTSPQFCDP*N
i	1	ļ	1			GQIRSYFTVLLRGLNETMLVK/PLCRREP/PEA
		1				GPGRQSTPAVTRDHRQHEDPRGAGRQWDAD
	1					PRPSAP/PAEVATGSRPGRHMWMRLCLAAQQ
}	}	1)		APGLPHRTSIRPGWRRLTEPEAWARRHRRPW
		1	1			GQRGAVRPPPQGAAPPPSHQGRRTNTDPSAT
1	1	1			}	PRLTVMSRCLAPDLKAPASGPRGWRRGMPQ
	1		1	1	ŀ	SS/GALLWTPPPTPRGSHSPRPREAPLRAIHPA
l		ı	1	1	Į.	GPSK/SRAGASGRLPEVIYGWVTLFTPPEAGT
		1	1	1	1	DA ADODONAL COD LA CICONDI MINERO COS DACOS DE
						F/LIPSPT*MSPALVIQPPVPPTQMGLRISGLPR QG*PSGAPW*LPGLAQLAFQCHLPHDEVGPP

PCT/US01/03800

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	""	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
00.,00		1	1 74.	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	Sequence	/=possible nucleotide deletion, \=possible
ŀ		1		sequence	i	nucleotide insertion
	 	1	 	Sequence		RNQSPLGNDTLSSGLPMGPRRQVWPLARVG
1		ĺ		ĺ		GHSSPREPQVLKKPLWGQTDIAGVGSASLYP
}]		1	ļ	DNL
362	1712	A	3136	1270	274	RVGMVLGTREVGDSTPPPSPPLYPFTGNEFVQ
""	****	''	3130	1270	214	HNTWQLSRVYPSDLRTDSSNYNPQELWNAG
	1					CQM/V*GGSRDWEEGVEEQQVGNKFSSDGR
	}	ļ	ļ	}		VGECSRKLLG*EMLSVDITSRYRAPSTYLLNS
			1			LKEGLEGLHGESCSSFLLGPSVAMNMQTAGL
						EMDICDGHFRQNGGCGYVLKPDFLRDIQSSF
		1		!		
		į				HPEKPISPFKAQTLLNQVISVQQLPKVDKTKE GSIVDPLVKVQIFGVRLDTARQETNYVENNG
		t d				FNPYWGQTLCFRVLGPDFPMLRFGKMDYDW
	ĺ			t		
i		(1			KSRNDLLGKTPCPGTCMQQGYRHIHLLSKDG ISLRPASIFVYICIQEGLEGDES
363	1713	C	3139	60	248	MFAGSYGKSMFSFSKKVLNCLPKWRYHFVIA
305	17.13	~	1337	00	240	PAMNESPLAPHLHQHLVFSVFQVLTILIGV**
364	1714	A	3140	57	418	SAFKTLQLPAFSLYFDLGSLKLLILRIHTSIVK
"	1 111	^	1 3140	"	710	NHKVESPRTMSPG*DPQSFLQIPQPRPPQLRV
	1			-		GLTSGLIQHFHSPSSCQFPLLRGPPFPRQPPLGI
	ļ	}	j .			SGASLCPVLSPPR*PLQPSSL
365	1715	A	3145	122	413	11 DADGE AFRICA LEGISOR
303	1713	^	3143	122	413	LLPYPSLFVFLRQCHFVT\RLECNGVVSAHCN LHLPGSSDSPASAS*VAGTTGVCHHTRLIF\VF
ļ		J				LV*TGFHYVAQAGLELLTA*S\PPQLPKVVGL QA
366	1716	A	3150	247	2	VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN
500	17.10	^	3130	247	2	FIKIKKLCIEGYY/NREPQNGRKIFANYVS\DK
		i i				GLMATIYEELLKLSNKLIQ
367	1717	A	3152	3	2367	QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT
	1,1,	**	3132	٦	2507	PIEKSDFAAATHPRAFYLSKPDETPNAWMSD
						SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS
		1	-			TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP
						NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN
						VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE
						SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS
						ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR
						QNLKEKHARHIADLRAYYESEINSLKQKLEA
						KEISGVEDWKITNQILVDRCGQLDSALHEATS
		}				RVRTLENKNNLLEIEVNDLRERFSAASSASKI
						LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF
(İ	ENAYKLSDDKEAQLKQENKMFQDLLGEYES
						LGKEHRRVKDALNTTENKLLDAYTQISDLKR
						MISKLEAQVKQVEHENMLSLRHNSRIHVRPS
						RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT
[1				QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST
	1					SLLIKKQRETSDTPIMRALKELDEGKIFKNWG
						TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK
]						CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR
						EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN
						ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV
					-	KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN
						DFEYTAKIRTLAETERFFDELTKEKDQIEAAL
					l	SRMPSPGGRITLQTRLNQVKCLSLNLL
368	1718	Α	3163	2 -	2350	EFKSGGCGAGLVAAGAVLVLYPASRAGERT
						RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL
					į	LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ
						MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT
-					·	TVPQTQGQTTAQKVSKKTGPRCSTALATGLK
			ł			NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP
L						AWDLKGQLCDLNAELKRCRERTQTLDQENQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						QLQDQLRDAQQQVKALGTERTTLEGHLAKV QAQAEQGQQELKNLRACVLELEERLSTQEGL VQELQKKQVELQEERRGLMSQLEEKERRLQT SEAALSSSQAEVASLRQETVAQAALLTEREER LHGLEMERRRLHNQLQELKGNIRVFCRVRPV LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLSRSD ERRGTLSGAPAPPTRHDFSFDRVFPPGSGQDE VFEEIAMLVQSALDGYPVCIFAYGQTGSGKTF TMEGGPGGDPQLEGLIPALRHLFSVAQELSG QGWTYSFVASYVEIYNETVRDLLATGTRKGQ GGECEIRRAGPGSEELTVTNARYVPVSCEKEV DALLHLARQNRAVARTAQNERSSRSHSVFQL QISGEHSSRGLQCGAPLSLVDLAGSERLDPGL ALGPGERERLRETQAINSSLSTLGLVIMALSN KESHVPYRNSKLTYLLQNSLGGSAKMLMFV NISPLEENVSESLNSLRFASKVEPSVLFGTAQS NRKWKTDPDLCVCVCVCVCVCVCVCVCVPMSMYRVRGGRVAGGCFIGWRAPCPRAIK
369	1719	A	3165	365	12	GYTSQGRWIDIERGPLTANTESLHENNFNALP GYIRKIE*1*IYKKN*INFGGVGLLNIVKISILS/K IYRFDAIPVKILTRFFINLDKLILKFVLKTKIAK NRIKTFYIMRRKKLGDSS
370	1720	Α.	3170	393	42	GASISPSAVIDGVEGLKPMQEQEAQEAGPCLD *HMAPEQWVAPR\RLLFRLIFSVLHALIIAAAA QSSAEEDEDPRN*GQSSEDQAPNQNGLIVIVH RVHVPLGAAATVPVHRSHFPR
371	1721	A	3173	770	510	GNGGCGLSQIPPSHLGAFSRGSLLSRG\DPRGP PPHPVIFFVFVVE\QGFTVLARMVSIS*PCDPP ALASQSAGITGVSHLARPQNLYF
372	1722	A	3180	381	76	RVLHHDNVPAHSSPQKREISQEFQLEIRHLP*S PDLAPSGCFLFLNLKNIFK\GTHFSLVDNVKK TVSTWLH/SQNAQFYKDRLNGWYHCLQKCL QHY*AYVEK
373	1723	A	3181	410	14101	RREVAGPEGKGLLLASAHTMLTPPLLLLPLL SALVAAAIDAPKTCSPKQFACRDQITCISKGW RCDGERDCPDGSDEAPEICPQSKAQRCQPNE HNCLGTELCVPMSRLCNGVQDCMDGSDEGP HCRELQGNCSRLGCQHHCVPTLDGPTCYCNS SFQLQADGKTCKDFDECSVYGTCSQLCTNTD GSFICGCVEGYLLQPDNRSCKAKNEPVDRPP VLLIANSQNILATYLSGAQVSTITPTSTRQTTA MDFSYANETVCWVHVGDSAAQTQLKCARM PGLKGFVDEHTINISLSLHHVEQMAIDWLTGN FYFVDDIDDRIFVCNRNGDTCVTLLDLELYNP KGIALDPAMGKVFFTDYGQIPKVERCDMDG QNRTKLVDSKIVFPHGITLDLVSRLVYWADA YLDYIEVVDYEGKGRQTIIQGILIEHLYGLTVF ENYLYATNSDNANAQQKTSVIRVNRFNSTEY QVVTRVDKGGALHIYHQRRQPRVRSHACEN DQYGKPGGCSDICLLANSHKARTCRCRSGFS LGSDGKSCKKPEHELFLVYGKGRPGIIRGMD MGAKVPDEHMIPIENLMNPRALDFHAETGFI YFADTTSYLIGRQKIDGTERETILKDGIHNVE GVAVDWMGDNLYWTDDGPKKTISVARLEK AAQTRKTLIEGKMTHPRAIVVDPLNGWMYW TDWEEDPKDSRRGRLERAWMDGSHRDIFVT SKTVLWPNGLSLDIPAGRLYWVDAFYDRIETI LLNGTDRKIVYEGPELNHAFGLCHHGNYLFW TEYRSGSVYRLERGVGGAPPTVTLLRSENPPI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				peptide	sequence	/=possible nucleotide deletion, \=possible nucleotide insertion FEIR\MYDAQHQQVGSNKCRVNNAGCSSLCL ATPGSRQCACAEDQVLDADGVTCLANPSYVP PPQCQPGEFACANSRCIQERWKCDGDNDCLD NSDEAPALCHQHTCPSDRFKCENNRCIPNRW LCDGDNDCGNSEDESNATCSARTCPPNQFSC ASGRCIPISWTCDLDDDCGDRSDESASCAYPT CFPLTQFTCNNGRCININWRCDNDNDCGDNS DEAGCSHSCSSTQFKCNSGRCIPEHWTCDGD NDCGDYSDETHANCTNQATRPPGGCHTDEF QCRLDGLCIPLRWRCDGDTDCMDSSDEKSCE GVTHVCDPSVKFGCKDSARCISKAWVCDGD NDCGDNSDEENCESLACRPPSHPCANNTSVC LPPDKLCDGNDDCGDGSDEGELCDQCSLNN GGCSHNCSVAPGGIVCSCPLGMELGPDNHT CQIQSYCAKHLKCSQKCDQNKFSVKCSCYEG WVLEPDGESCRSLDPFKPFIIFSNRHEIRRIDLH KGDYSVLVPGLRNTIALDFHLSQSALYWTDV VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT TLLAGDIEHPRAIALDPRDGILFWTDWDASLP RIEAASMSGAGRRTVHRETGSGGWPNGLTV DYLEKRILWIDARSDAIYSARYDGSGHMEVL RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA KANKWTGHNVTVVQRTNTQPFDLQVYHPSR QPMAPNPCEANGGQGPCSHLCLINYNRTVSC ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR GVDLDAPYNYIISFTVPDIDNVTVLDYDARE QRVYWSDVRTQAIKRAFINGTGVETVVSADL PNAHGLAVDWVSRNLFWTSYDTNKKQINVA RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY WTDGDNISMANMDGSNRTLLFSGQKGPVGL AIDFPESKLYWISSGNHTINRCNLDGSGLEVID AMRSQLGKATALAIMGDKLWWADQVSEKM GTCSKADGSGSVVLRNSTTLVMHMKVYDESI QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC MCTAGYSLRSGQVACEGVGSFLLYSVHEGIR GPLDPNDKSDALVPVSGTSLAVGIDFHAEND TIYWVDMGLSTISRAKRDQTWREDVVTNGIG RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG SFRYVVISQGLDKPRAITVHPEKGYLFWTEW GQYPRIERSRLDGTERVVLNVSISWPNGISV DYQDGKLYWCDARTDKIERIDLETGENREVV LSSNNMDMFSVSVFEDFTYWSDRTHANGSIK RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR QKGTNVCAVANGGCQQLCLYRGRGQRACA CAHGMLAEDGASCREYAGYLLYSERTILKSI HLSDERNLNAPVQPFEDPEHMKNVIALAFDY RAGTSPGTPNRIFFSDIHFGNIQQUNDDGSRRIT IVENVGSVEGLAYHRGWDTLYWTSYTTSTTT RHTVDQTRPGAFERETVITMSGDDHPRAFVL
						DECQNLMFWTNWNEQHPSIMRAALSGANVL TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF WTDWVRRAVQRANKHVGSNMKLLRVDIPQ QPMGIIAVANDTNSCELSPCRINNGGCQDLCL LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR AQDEFECANGECINFSLTCDGVPHCKDKSDE KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN GADDCGDGSDEIPCNKTACGVGEFRCRDGTC IGNSSRCNQFVDCEDASDEMNCSATDCSSYF

NOC of nucleotide order of the control of the contr	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence uence	1 -				beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
uence 09496 007890ndi 1914 ng to first amino acid residue of peptide residue of peptide sequence 0-Gilutamine, R-Arginine, S-Serine, S-Serine, peptide sequence 0-Gilutamine, R-Arginine, S-Serine, S-Serine, peptide sequence 0-Gilutamine, R-Arginine, S-Serine, S-Serine, periodic sequence 0-Gilutamine, R-Arginine, S-Serine, S-Serine, periodic sequence 0-Gilutamine, R-Arginine, S-Serine, sequence 0-Gilutamine, R-Arginine, R-Arginine, S-Serine, sequence 0-Gilutamine, R-Arginine, R-Arginine, R-Serine, sequence 0-Gilutamine, R-Arginine, R-Arginine, R-Arginine, R-Arginine, R-Arginine, R-Serine, R-Arginine, R-Arginine, R-Britamine, R-Arginine, R-Britamine, R-Arginine, R-Britamine, R-Arginine, R-Britamine, R-Arginine, R-Britamine, R-Arginine, R-Britamine, R-Arginine, R-Britamine, R-Arginine, R-Britamine, R-Arginine, R-Britamine, R-Arginine, R-Britamine, R-Arginine, R-Britamine, R-Britamine, R-Britamine, R-Britamine, R-Britamine, R-Britamine, R-Britamine, R-Britamine, R-Britamine, R-	nucl-	peptide	f	in	nucleotide	location	
pepide residue of pepide sequence pepide seque	eotide	seq-		USSN	location		
mino acid residue of peptide sequence peptide sequence sequence sequence peptide sequence per sequence per sequence sequ	seq-	uence					
residue of peptide pep	uence		ļ	914			Q=Glutamine, R=Arginine, S=Serine,
peptide sequence Including insertion							T=Threonine, V=Valine, W=Tryptophan,
Inclodide insertion			ł			sequence	
RIGVKGVLFQCERTSLCYAPSWCDGAND CGDYSDERDCPGVKRPCCLNYRCSGACIP MSWTCDKEDDCEHGEDETHCNKFCSEAQPE CONHRCISKQWLCDGSDDCGDGSBCAAHCE GKTCGPSSPSCPGTHVCVPERWLCDGDKDCA DGADESIAGCLVNTCDDREFMGVRQCIP KHFVCDHDRDCADGSDSEPECYPTCGPSEF RCANGRCISSRQWLCDGHENCENDCDDSDEAPK NPHCTSPEHKCNASSGFLCSSGCVARALLCN GDDDCGDSSDERGCHNECLSRLSGCSQDC EDLKIGFKCRCRPGFRLKDDGTTCADVDECS TIFFCSGCNITHGSSYKLCLVGGVAPRGGDP HSCKAVTDEEPFLIFANRYVLRKINLDGSNY TLLKQGLNNAVALDPDTREGMTYMDVTTO GSMURNHLINGSNVQVHLRTGLSNPDGLAV DWVGGNLWCXGRGDDTTSKSKLNGAVARVIV VSSGLREPFALVDDVRGOMTYMDVTTO GSMURNHLINGSNVQVHLRTGLSNPDGLAV VSSGLREPFALVDVQNQVLYNTDWGDMSL TRIPAGNASSGNYDVHTTDWGANGAYRVL VSSGLREPFALVDVQNQVLYNTDWGDMSL VSSGLREPFALVDVQNQVTAVTDWGDMSL VSSGLREPFALVDVQNQVTAVTDWGDMSL VSSGLREPFALVDVQNQVTAVTDWGDMSL VSSGLREPFALVDVQNQVTAVTDWGDMSL VSSGLREPFALVDVGNGCSSNVVVLSQDIPH HFALTLEDVYVYTDWGTSTSNRAKKTTGTN KTLLISTLHRPMDLHVFHALRQPDVPNNFCX VNNGCSSLLLSRGGGRKACPTNVLGSD GKTCVSNCTASGPVCKNDKCDFWWKCDTE DDCGDHSDFPFDCFFKCRCPQCCSTGLCTN PAFCGGDNCQDNSDEANCDHVLSQDPP FORSTER CTNTNRCPGGFRCNGQDNCGDEDERLPCPP VTCAPMQPQCSTKRCPPRWCCDEDERLPCPP VTCAPMQPQCSTKRCPPRWCCDEDERLPCPP VTCAPMQPQCSTKRCPPRWCCDEDERLPCPP VTCAPMQPQCSTKRCPPRWCCDEDERLPCPP TCASSGRAVCGGGGCGGCGGCGGCGGCGGCGGCGGCGGGGGGGGGG	İ	İ	ĺ				
CGDYSBERDÉPOVREPRÉCALVITACESGACIP MSWTCDKEDDCEHGEDETHONKESAQPE CQNHECISKQWL_CDGSDDCCDGSDEAAHCE GRTCGPSSSECPTHVVPERWLCDGDKDCA DGADESIAAGCL_YNSTCDDREFMCQNRQCIP KIFVCDHEDRDCADGSDESPECEYPTCGPSEE RCANGRCLSSRQWECDGENDCHDQSDEAPKE NPHCTSPEHKCNASSOFLESSGRCVAEALLCN GQDDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGGKCRCROPERLEDDCRTCAAULCN GQDDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGGKCRCROPERLEDDCRTCAAULCN GQDDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGGKCRCROPERLEDDCRTCAAULCN GQDDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGGKCRCROPERLEDDCRTCAAULCN GODDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGGKCRCPORTLEDDCRTCAAULCN GODDCGDSSERGCHINECLSRKLSGCSQDC EDLKIGGKCRCPORTLEDDCRTCAAULCN GODDCGDSSERGCHINECLSRKLSGCSQDC EDLKIGGKCRCPORTLEDDCRTCAAULCN GODDCGDSSERGCHINECLSRKLSGCSQDC EDLKIGGKCRCPORTLEDDCRTCAAULCN GODGCMSCRCCAAULCN GODGCMSCRCCAAULCN GODGCMSCRCCAAULCN GODGCMSCRCCAAULCN GODGCMSCRCCAAULCN GODGCMSCRCCAAULCN GODGCMSCRCCAAULCN GODGCMSCRCCAAULCN GODGCMSCRCCAAULCN GOOGCMSCRCCAAULCN GODGCCAAULCN GOD			ļ		sequence		
MSWTCDKEDDCEHGEDETHICNIFCSEA,OPE CONHRICISKOWLOGSDEAGAHCE GKTCGPSSTSCGTHVCVPERWLCDGNECA DGADESIAAGCL VINSTCDDREFNCONROCCIP KHPVCDHIDDCADGSDESIPECEYPTCGFSEE RCANGRCLSSRQWCDGENDCHDQSDEAPK NPPICTSPEHKCNASSQFLCSSGRCVABALLCN GQDDCGDSSDERGCHINECLSRGCODC EDLKIGFKCRCRPGFRLKDDGRTCADVDECS TTFPCSGRCTHTGSYKELCVEGYAPEGGDP HSCKAVTDEEPFLIFANRYVLKLNLDGSNY TLLKOGLNNAVALDFDYREQMFWTDVTTTQ GSMERRMHLNGSNVQVLHRTGLSNPDGLAV DWVGGRLYWCDKGRDTIEVSKLNGAYRTVL VSGGLREPRALVVDVQNQYLWTDWGDHSL ICRIGMBGSSSRSVVDTSKTWPRGLTDVTTE RIVWADABEDVIEFASLDGSNRHVVLSQDIPH FALTLEEDYVYWTDWTKISSRAKTITDVTTE RIVWADABEDVIEFASLDGSNRHVVLSQDIPH FALTLEEDYVYWTDWTKISSRAKTITGVT KTLLISTLIRPRADLEVPHALROPDVPNEPCK VNNGCSNLCLLSPGGGHKCACPSWWCCDTE DCGODHSDEPDOPEPFKCRPGOPCSTICTIN FARIDGENDCQDNSDEANCDHIVCLPSQC VNNGCSNLCLLSPGGGHKCACPSWWCCDTE DCGODHSDEPDOPEPFKCRPGOPCSTICTIN PARIDGENDCQDNSDEANCDHIVCLPSQC VTCAPNOPCSTITCRCTPRVWCDDEDDCVD GSDEPANCTQATCGVDFFRCDSGRCCPARW KCDGEDDGGDSSEPPEDCEDERTCPS VTCAPNOPCSTITCRCTPRVWQDEDDDCVD GSDEPANCTQATCGVDFFRCDSGRCCPARW KCDGEDDGGDSSEPPECCDERTCPSQC VTCAPNOPCSTITCRCTPRVWQDEDDDCVD GSDEPANCTQATCGATCGATCCACC KNNRCVPGRWQCDTDNDCDNSDESCTRE PCSESSEPSCANGCGARWKCDGDIDCADGS DEDCTRRCDDPCCDNSDEANCDHIVCLPSQC KNNRCVPGRWQCDTDNDCDNSDESCTRE PCSESSEPSCANGCGARWKCDGDIDCADGS DEDCTRRCDMDPQCKSGHCTLRWRCD DADACMBGDEEACCTGATCNTCPLDFQCNNT LCCPLAWKCDGEDDCGDNSDENSERECCACC VPNRPRCCNDRVCLWGRQCDDIDCADGS DECCTRRCDMDPQCKSGHCTLRWRCD DADACMGDSEEACCTGATCTCDFTQPTQNNT LCCPLAWKCDGEDDCGDNSDENPECARRV CPPNRPRCCNDRVCLWGRQCDDTDCGD GTDEEDCEPFTAHTTHTCKGRCEFLCRNQRC CPNRPPRCCNDRVCLWGRQCDDFKLTSCAT NASICGBEARCVRTSKAAVCACRSGFHTVYG QPGCQDNSCLARGCGTGCSQLCNNTKGGHLCSC ARREMETINTCKAEGSERECTWIK RQDGGVTHLNISGLKAPRGIADDWAGNNY WTDSGRDVEVAQMKGGNRKTLISGMDPBFD AKPMGTNTNTKGGHLCSC ARREMETINTCKAEGSERECTWIK RQDGGVTHLNISGLKAPRGIADDWAGNNY WTDSGRDVEVAQMKGGNRKTLISGMCDPFTP PDAPRGTCNLCCPRGGSCCTAARRGPKCCP PPNTGCTVPNGRCLONGCTCAAPSGR MPTCRCPGTTGFTCTQVCAGCSCGTAATSCMRACCHASTCT VNQGNQPQCRCLARGCGTCAAPSGR MPTCRCPGTTGGCAGCSCTANSTCT VNQGNQPCCCLGPGGGSCCTAATSCMRACCCC PPNTGTCQWAADGSGCTCAATSGRGCCCTAATSCARGCCCCAATSCGCCCCCC PROTCQWAADGSCGTAM							
CONHRCISKOWLEDGSDEAGEGSEAEACE GKTCOPSSTSCE/GTIVCVEPRAWLCGDKDECA DGADESIAAGCL_YNSTCDDREFMCONROCIP KHEVCOHNDRO_DAGSDESPECE/PTGGPSEE RCANGRCI_SSRQWECDGEDCHDQSDEAF WHEVCTSPEHKCHASQQL_CSSRCV_AEALLCN MCDDCGDSSDERGCHENGCLSRKLSGCSQDC EDLKIGFKCRCREPGFILKDDGRT_CANDECS TITFPCSQRCINTHGSYKLLOVEGYAPRGGDP HSCKAVTDEEPFLIFANRYYRKLNLDGSDY TLLKQGLNNAVALDFDYREQMITYTDVTTQ GSMRRMHILMGSNVQVLHRTGLSNPDGLAV DWVGGRLYWCDKROTETEVSKLNGAYRTVL VSSGLREFRALVVDVQNGYLWYTDWGFSISI IGRIGMGSSSRSVYUTSKTIPWHOLILDYVTE RIYWADABEDYIEFASLDGSNRHVVLSQDIPH FIALTLFEDYYYWTDWGTSSINAKHKTGTN KTLLISTLHRPMDLHVFHALRQPDYNHIPECK VNNGGCSNICLLSPGGGHKCACFUNYLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGGDHSDEPPDCEPFKCRPGQFCSTGICTN PAFICDGDNDCQDNSDEAANCDHYCLSGKK CTNTNRCRGFFRGNGQDNCCGBEBEDCFE VTCAPNQFQCSITKRCIPRYWCDRDNDCVD GSDEFANCTQMTCOVDERCKSDRGCPARW KCDGEDDCDGDSDEPKEECDERTCEPYOFR CTAPNQFQCSTTRCCIPRYWCDRDNDCVD GSDEFANCTQMTCOVDERCKSDRGCPARW KCDGEDDCDGDSDEPKEECDERTCEPYOFR CSSEFSCANGRCLAGRWKCDGDHDCADD GDDCDCTPRCMDDQFQCKSGHCIFRWRCDB DEADCTPRCMDDQFQCKSGHCIFRWCDB DEADCMGSSDERACGTGTTCPLDEFQCNNI LCPLAWKCDGGDDCGONSDEESCTPR PCSESEFSCANGRCLAGRWKCDGDHDCADD DADCMGGSDEACEGTGTTCPLDEFQCNNI LCPLAWKCDGGDDCGNSDEESCTPR PCSESEFSCANGRCLAGRWKCDGDHDCADD DADCMGGSDEACEGTGTTTTNSRHR RCGDBACVTTHTHCKLKSEFLCRNORCL SSSLRCMMFDCGGSSVGLDAMDVHVKA GRYVWTNWHTGTSYRSLPAAGCAGCTTNSRH RCGDBACVTTHTHCKLKSEFLCRNORCL SSSLRCMMFDCGGSSVGLDAMDVHVKA GRYVWTNWHTGTSYRSLPAAGCAGGTTNSCHAR RCGDBACVTTHTHCKARGSEFYOU YANDNEERS LPFGHPHSAYEQAFQGDESVRIDAMDVHVKA GRYVWTNWHTGTSYRSLPAAGCAGFHTVPD ANDROGDATHTHNSGLKAPRGIADDWAGNVY WTDSGDATEVAGRACGGTTNCTRAGHLCS CARREST LPGGTMYNDFGCRGCCGCTNCTCG CSSLRCMMFTTHYTCKARGGCTTNSCHAR RCGDBACVTTNSTRAGHACTACAAA RLSVGGGACCTTATYCGTRAGCCC CSSLRCMMFTTHYTCKARGGCTCATASTSO UPTCCTTTTTTTNSTRH RCGNCTATASTSCAT NASICGDBACVTTNTTNGHHACTACAAAA RLSVGGGCACATATCPTAGGACCAASPSG MPTCCCCCAACAASTCT VNGGNGPCCCACAACAASTCT VNGGNGPCCCACAACAASTCT VNGGNGPCCCACAACAASTCT VNGGNGPCCCACAACAASTCT VNGGNGPCCCACAACAASTCT VNGGNGPCCCACAACAASTCT VNGGNGPCCCACAACAASTCT VNGGNGPCCCACAACAACAACAACAACAACAACAACAACAACAACA							
GKTCGPSSTSCGTHVCVPERWLCDGMEAC DGADEMAAGCL VINSTCDDREEMCONRCCIP KHPVCDHINDCADGSDESPECEYPTCGPSEE RCANGRCLSSRQWCDGEDDCIDGSDEAPK MPHCTSPPHKCNASSQFLCSSGRCVASALLC GQDDCGDSSDERGCHINECLSKLSGCSQDC EDLKIGFKCRCRPGFRLKDGRTCADVDECS TTPFCSQRCINTHGSYKCLCVEGYAFRGGDP HSCKAVTDESPFLLFANRYTIRKINLDGSNY TLLKQGLNNAVALDFDYREQMTYWTDVTTQ GSMREMBLIAGSNVQVLHETGLSSPDGLAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTYL VSSGREPRALVDVQNGYLYWTDWGDHSL IGRIGMGGSSRSVIVDTKITWPNGLTLDYVTL VSSGREPRALVDVQNGYLYWTDWGDHSL IGRIGMGSSRSVIVDTKITWPNGLTLDYVTL TRIYWADAEDYIEFASLDGSNRHVVLSQDIPH IFALTLFEDYYWTDWETKSINRAHKTTGTN KILLISTLIHRMDLIVFHALRQPDYPNHPCK VNNGGSNLCLSPGGGHKCACTPTSYLGS GRTCVSNCTASGPVCKNDKCTPFWKCDTE DDCGDHSDEPPDCPEFKCRPGQPCCSTGICTN FAPICOGDNCQDNSDEANCDHVCLPSQFK CTNTRCCHGFRCNGQDNCGGGEDERDCPE VTCAPNQPCSTITKCPPVWCDBDNDCVD GSDEFANCTOMTCGVDEFRCKDSGRCPARW KCDGEDCGDOBSPFRECDESTCEPYOFRC KCNRCVPGRWQCDYDNDCGDNSDESCTPR PCSSESFSCANGRCLAGRWKCDGDHDCADG SDEADCTRRCMDQFQCKSGCTPLRWRCDA DADCMGSDEBACGTGRTCPHOFRC KNRCVPGRWQCDYDNDCGDNSDESCTPR PCSSESFSCANGRCLAGRWKCDGDHDCADG GTDEEDCEPPTAHTTHCKDKKETLCANGC SSSLRCMMFDDCGGGSBECRICLER CPPNEPRCKNDRVCLWGRQCGTDNCGG GTDEEDCEPPTAHTTHCKDKKETLCRNQRCL SSSLRCMMFDDCGGGSSVRIDAMDVITVA GROEGORIC CRAFTITHCKDKKETLCRNQRCL SSSLRCMMFDDCGGGSSVRIDAMDVITVA GROEGORIC CRAFTITHCKDKKETLCRNQRCL SSSLRCMMFDDCGGGSSVRIDAMDVITVA GROEGORIC CRAFTITHCKDKKETLCRNQRCL SSSLRCMMFDDCGGGSSVRIDAMDVITVA GROEGORIC CRAFTITHCKDKKETLCRNQRCL SSSLRCMMFDDCGGGSSVRIDAMDVITVA GROEGORIC CRAFTITHCKDKKETLCRNQRCL SSSLRCMMFDGCGGGRAVGCGGTNCGGL LERTLVQDNQWTTINGGHLCSC ARNYMTHMTGTSTYSLIPPAPTTNSNHA RQDBGVTHLNISGLKMPGGLADDVATCHCDWKCC CLSSSCRCMFDGCGGCGCTAARRQPKCC PPCYTGTMYTHLNISGLKMPGGLADDVATCHCTONACH CLSSSCRCMFTCGGGCCTAARRQPKCC PPCYTGDKCGLDCGCGCTAARRQPKCC PPCYTGDKCGLDCGCCTAARRQPKCC PPCYTGDKCCLDCGCCTAARRQPKCCC PPCYTGDKCCELDCCCHARCCGCCAARROPCCC PPCYTGDKCCLDCCCCC NGCTCQMAADGSCGCTAAYFGGGCCTAARROPCC PPFPPP PDAARPGTCNLCCFNGGCCTAARROPCCCPPPH ATVOPPLGGTMYTGCLGCTCAARROPCCCPPPH TORQMCPCLCCCTAGGCCTAARROPCCCPPPH TORQMCPCCLGGGCCTAARROPCCCPPPH TORQMCPCCLGGGCCTAARROPCCCPPPH TORQMCPCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	ì]		,		
DGADESIAAGCL_YNSTCDDREFMCQNRCQIE KHFVCDHRDRCDAGSDESPECELPYTGGFSEF RCANGRCI_SSRQWECDGENDCEDQSDEAF NPHCTSPEHRCNASSQFL_SSRGCVASALLCN GQDDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGFKCRCRPGFHLKDDGRTCADVDECS TTFPCSQRCINTHOST SYKLCLVEGYAFRGGDP HSCKAVTDEEPFLIFANRYTLRKLNDGSRTY TLLKQGLNNAVALDFDYREQMITYTDVTTQ GSMIRRMHLNGSNVQVLHRTGLSNPDGLAV DWVGGNL_YWCDKGRDTIEVSKLNGAYRTVL VSSGLREFRALVVDVQNGYL_YWTDWGDHSL IGRIGMGSSRSVYDVDTKTWPNGHLTLDYVTE RIYWADAREDYIEFASLDGSNR;NVLSQPIPH PRICTIFEDYVYWTDWTHSSRNAKHTIGTIN KTLLISTLIBRMDLHVHHALRQPDVPNHPCK VNNGGCSNLCLLSPGGHKCACPTNYLGSD GRTCVSNCTASQPVCKNDKCPPWWCDTF DDCGDDSDEPPDCPFEKCRPQCSTGICTIN PAFICOGDNDCQDNSDEANCDHVCLPSQFK CTNTRRCHGFRCNGQDNCGDGEDERDCPE VTCARNQPQCSTRTKCPRVWCODRDCVD GSDEPANCTQNTCGVPERCKBGCGPARW KCDGEDDGODSDEPPDECERFCRYQCSTGICTON PAFICOGDNCGDDGDEDBRDCPE VTCARNQPQCSTRTKCPRVWVCORDNDCVD GSDEPANCTQNTCGVPERCKBGCCPARW KCDGEDDGODSDEPPDECERFCRQCTGNCTOL GSDEPANCTQNTCGVPERCKBGCCPARW CDGDBGDDGDGDSDEPPECCBRTCQDDGDGDDGDDCDD GDNSDEPDCGDNSDESSCTPR PCSESFSCANGRCAGRWKCDGDBCADG DEDCCPPRCCMGGDGGDGDGDCADG DEDCCPPTAHTTHCKDKEFLCAVGCDN LCPLAWKCDGGDDCGDNSDENFECCARFV CPNNRPFRCKNDRVCLWIGRGCDTDCGD GTDEDCCPPTAHTTHCKDKEFLCAVGTNCG LSSLRCAMGDDCGDSSDENFECCARFV CPNRPFRCKNDRVCLWIGRGCDTDCGD GTDEDCCPPTAHTTHCKDKEFLCAVA SSLRCAMGDDCGDSSDENFECCARFV CPNRPFRCKNDRVCLWIGRGCDTDCGD GTDEDCCPPTAHTTHCKDKEFLCAVAGGHLCSC SSLRCAMGDDCGDSSDESVEIDAANDFIXAA RIGGGDBARCVRTEKAAVCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARPFMRTINTCKABGSETQVLYIADDNEIRS LPFGIPHSAYEQACPQODBSVRIDAANDFIXAR RQDRGVTHLNISGLKAPRGIADWAGNNCH LSVIGSRILNGTDPIYAADSKRCLSHPPSIDV FEDYYGVTYNNTGYTGITYSRSLPFAAPFTISNTRIR RQDRGVTHLNISGLKAPRGIADWAGNNC LSVIGSRILNGTDPIYAADSKRCLSHPPSIDV FEDYYGVTYNNTGYTGITYSRSLPFAAPFTISNTRIR RQDRGVTHLNISGLKAPRGIADWAGNNC LLSTGGRGCTLAYRGCSCCAPC QPRYTODKCELLOQUCCWECKRINGGCAASSRG WPTCCCTGTTQPKCTQQCCGTCAASSRG WPTCCCTGTTQPKCTQQCCGTCAASSRG WPTCCCTGTTQPKCTQQCCGTCAASSRG WPTCCCTGTTQPKCTQQCCGTCAASSRG WPTCCCTGTTQPKCTQQCCGTCAASSRG WPTCCCTGTTQPKCTQQACGTCAANSTCT VNQGNQPQCCLAGGGCTLAYRGCCGTCAASSRG WPTCCCTGTTQMADGGSCTLAYRGCCCTCAPFINGTCAASSRG WPTCCCTGTGAACSCTAASGCCTCATCTGGRCC	1						
KHPVCDHIRDCADGSDESPECEYPTICGFSER RCANGRCLSSRQWECDGENDCHDQSDEAPK NPHCTSPEHKCNASSQFLCSSGRCVASALLOW GODDCODSSERGCHNECLSKLSGCSQDC EDLKIGFKCRCRPGFILKDGRTCADVDECS TITPCSQRCINTHGSYKLCLVEGVAPRGGDP HSCKAVTDEEPFLIFANRYYLRKLNLDGSNY TLLKQGLNNAVALDFDYREQMTWTDVTYTQ GSMRRMHLMGSNVQVLHRGLSIPPOLAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTYTU VSSGLREPRALVDVQNGYLYWTDWGDHSL IGRIGMGGSSRSVIVDTKITWPNGLTLDYVTY VSSGLREPRALVDVQNGYLYWTDWGDHSL IGRIGMGSSRSVIVDTKITWPNGLTLDYVTTY TILLTHERDUNGWGNGTHEVSKLNGAYRTYTU VSSGLREPRALVDVQNGYLYWTDWGDHSL IGRIGMGSSRSVIVDTKITWPNGLTLDYVTTY THE RIYWADAEDYTERSJLGSNRHVVLSQDIPH IFALTLFEDYYWTDWFTKSNRAHKTTGTN KITLLISTLHRWBLDHVFHALRQPDYPNHPCK VNNGGSNLCLSPGGGHKCACPTNYYLGSIS GRTCVSNCTASGPVCKNDKCIPPWWCDTE DDCGDHSDEPPOCPEFKCRPGGPGCSTGICTN PAFICIODNOCQDNSDEANCIPHVCLPSQFK CTNTRRCRGFRCNGQDNCGDGEDERDCPE VTCARNQPCGSTIKTCPPWWCDDFD PAFICODDNOCQDNSDEANCIPHVCLPSQFK RCTNTRRCRGFRCNGGDNCGDGEDERDCPE VTCARNQPCGSTIKTCPPWWCDDDDCVD GSDEFANCTQMTCQVDFRCKDSGRCPAW KCDGEDCGDGSPEPRECDEFT PAFICODDNOCQDNSDEANCIPHVCLPSQFK RCTNTRRCRGFRCNGGDNCGDGEDERDCPE VTCARNQPCGSTIKTCPPWWCDDDDCVD GSDEFANCTQMTCQVDFRCKDSGRCPAW KCDGEDCGDGSDEPRECDERTCEPYQFRC KNRCVPGRWQCDYDNDCGDNSDESCTGPYQFRC KNRCVPGRWQCDYDNDCGDNSDESCTGPYQFRC KNRCVPGRWQCDYDNDCGDNSDESCTGPYQFRC KNRCVPGRWQCDYDNDCGDNSDESCTGPYQFRC KNRCVPGRWQCDYDNDCGDNSDESCTGPYQFRC CNRCVPGRWQCDYDNDCGDNSDESCTGPYQFRC KNRCVPGRWQCDYDNDCGDNSDESCTGPYQFRC KNRCVPGRWQCDYDNDCGDNSDESCTGPYQFRC CNRCVPGRWQCDYDNCGDNSDESCTGPYQFRC KNRCVPGRWQCDYDNDCGDNSDESCTGPYQFRC CNRCVPGRWQCDYDNCGDNSDESCTGPYQFRC CNRCVPGRWQCDYDNCGDNSDESCTGPYQFRC CNRCVPGRWQCDYNDCGDNSDESCTGPYQFRC CNRCVPGRWQCGNCGRCGTCAPTFTCGNC SSSLRCMMFDDCGGSSVRIDAMDVTYNAV WTDSGGDFANCTGRAFTTHCKCKREFTLANGATA GRYYMTWHITGTSYRSLPPAPTTNSRHA RQDGGVTHLNISGLKMPGGLADPHYTDGPKKCCC PREYTONCGELDCCWRRCGNGCTCANFSTCT VNQCRQCPCCLAFGGGCTLANGCCCANNSTCT VNQCRQCPCCLAFGGGCTLANGCCCANNSTCT VNQCRQCPCCANGSCGCTANSTCTMCCCC PRYTONCGELDCCWRGAGCGTCANFSTCT VNQCRQCPCCANGSCGTCANSTCTCMCCCC PRYTONCGELDCCWRGAGCGTCANFSTCT VNQCRQCCCANGSCGCTCANFSCROPCCC PREGSGCCTANFSCMMPCCCCPPHM TGPRCEBETVSQQQCGHAASILPLLLLLLVL VAUVVPWYKRYCQGAGGGTQCASTCFSRCVM							
RCANGRCLSSRQWECDGENDCHDQSDEAPK NPHOTSPEHKCNASSQFLCSSRCVABALLCN GODDCGDSSDERGCHINECLSKLLSGCSQDC EDLKJGFKCRCRPGFLKDDGRTCADVDECS TTPCSQRCINTHGSYKCLCVEGYAPRGGDP HSCKAVTDEEPELLARNRYLJKLNDGSNY TLKQGLNNAVALDFDYREQMIYWTDVTTQ GSMIRRMHLNGSNYQVLHRTGLSNPFDGLAV DWVGONLYWCDKGRDTIEVSKLNGAYRTVL VSSGREPRALVVDVQNGYLYWTDWGDSL IGRIGMDGSSRSVIVDTICTIWPHGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALTLFEDYVYTDWETISSRNAHKTTGTN KTLLISTLHPPMDLHVFHALRQPDVPHHPCK VNNGCSNLCLISPGGGHKCACPTNPYLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPOCPETKCRPGQFQCSTGICTN PAFICCGDNDCQDNSDEANCDHVCLPSQFK CTNTNRCPGEFRCNGQDNCGDEDEDCDE VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDEFANCTQMTCGVDEFRCKDSGRCTPARW KCDGEDDCCDGSDEFRCEDEPCEPK KNRCVYGRWQCDYDNDCCDNSDESCTPR PCSESEFSCANGRCIAGRWKCDGBEDCADGS DEKDCTTRCDMDQFQCKSGHCIPLRWRCDA DADCMGSDSEAACGTGVRTLGCDFTQCNNT LCKPLAWKCDGEDDCGDNSDENFECARFV CPPNFFRCKNBRVCLWIRGQCDGTDNCGD GTDEEDCEPFTAHTTHCKDKKEFLCRNORCL SSSLACNMFDDCGGOSDENKELDEFQCNNT LCKPLAWKCDGEDDCGDNSDENFECARFV CPPNFFRCKNBRVCLWIRGGCDGTDNCGD GTDEEDCEPFTAHTTHCKDKKEFLCRNORCL SSSLACNMFDDCGGOSDENKELDGRNCH SSSLACNMFDDCGGOSDENKELSGRNCH ANSICGDBACCVRTEKAAYCACRSGFHTVPG QPGCODNECLRETGCTSGCJCNTKGGHLCSS ARNFMKTTHTCKAEGSEYQVLYIADDNERS LPFGHPBAYEQAPQGDESVRIDAMDVHVAA GRVWTNWHTGTISTRSLIPFAAPFTTSNRHR RQDBRCVTHINSIGLKRWRGAIDWVAGNNY WTDSSGRDVEWAQMKGGENKTLISGMDEPH AIVVPLRGTMYWSDWGNPKETAAMGT LRETTLYQDNIQWFTGLAVDYNBRELYWADA KLSVIGSIRLNGTDFTVAADSKRGLSHFFSIDV FEDYTYGTYTJNNRVFKLHKERPLVVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKCCEWL CLLSPGPVCTCPNGRLIDAGTCAASPSG MPTCRCPTGFTGRCCTPGGSCTLNARRQDYCRC QPRYTGDKCELTQNGRCCAASPSG MPTCRCPTGFTGRCCTPGGSCTLNARRADGT LRETTLYQDNIQWFTGLAVDYNBRELYWADA KLSVIGSIRLNGTDFTVAADSKRGLSHFFSIDV FEDYTYGTYTNNRVFKHKREPRLVVLTLTGG LSHASDVVLYHQHKQPEVTNPCDRKCCEWL CLLSPGGPCCTPGGRCCTHAGRTCAASPSG MPTCRCPTGFTGFRCTQQVCAYCANNSTCT VNQGNQPCCRCLPGGTGGGCTHARRADGT ARTTHOTOMAGDSRCCRCTAPEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTIGRTCAAPSG MPTCRCPTGFTGFRCTONGCTCAAPSG MPTCRCPTGFTGFRCTONGCTCAASPG MPTCRCPTGFTGFRCTONGCTCAASPG MPTCRCPTGFTGFRCTONGCTCAASPG MPTCRCPTGFTGRCTONGCTCAASPG MPTCRCPTGFTGRCTONGCTCAAPSG	ŀ						
GODDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGFKCRCRGFFLKDRCADVDECS TTPCCSRCINTHGSYKCLCVEGYAPRGGDP HSCKAVTDEEPFLJEANRYVJRKLMLDGSNY TLLKQGLNNAVALDFDYREQMITWTDVTO GSMIRRMHLNGSNYQULHRTGLSNPDGLAV DWVGGNLYWCDKGRDTDEVSKLNGAYRTVL VSSGLREFRALVVDVONGYLYWTDWGDHSL IGRIGMDGSSSRVJOTNTGTWFRGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALTLEEDYVYMTDWETISSNRAHKTTGTN KTLLISTLHRPMDLHVFHALRQPDVPNHPCK VNNGGCSNLCLISPGGGHKCACTFNYTLGSDIPH IFALTLEPDYVYMTDWETISSNRAHKTTGTN KTLLISTLHRPMDLHVFHALRQPDVPNHPCK VNNGGCSNLCLISPGGGHKCACTFNYTLGSDIPH IFALTLEPDYVYMTDWETISSNRAHKTTGTN KTLLISTLHRPMDLHVFHALRQPDVPNHPCK VNNGGCSNLCLISPGGGHKCACTFNYTLGSGF GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGGHSDEPPDCEFFKCRPGPQCSTGICTN PAFICDGDNDCQDNSDEANCDHVCLPSQFK CTNTNSCPIGFRCKQDNDCGGEDERDCPE YTCAPNQFQCSITKRCPRVWVCDRDNDCVD GSDEPANCTQMTGTWDEFRKCSGRCPARW KCDGEDDCGDDGSDERECCDERTCEPYQFRC KNNRCVPGRWQCDYDNDCCDDNSDESSCTRR PCSESEFSCANGRCLAGRWKCDGBEDCADGS DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMGSDESEACGTGVRTCPLDEFPCNNT LCKPLAWKCDGEDDCGDNSDESCTRR PCSESEFSCANGRCLAGRWKCDGBEDCARDS DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMGSDESEACGTGVRTCPLDEFPCNNT LCKPLAWKCDGEDDCGDNSDENECARPV CPPNRFFRCKNBRVCLWINGQCDGTDNCOD GTDEEDCEPFTAHTTHCKNKKEFLCRNNCL SSLLKCNMFDDCGDGSDEDESUPKLTSCAT NASIGGDBACVRTEKAAYCACRSGFHTVPG QPGCDDNECLREGTCSQLVATIAGDHLCS CREERGTSQLVATIAGGHLCSC CREERGTSQLVATIAGGHLCSC CREERGTSQLVATIAGGHLCSC CREERGTSQLVATIAGGHLCSC CREERGTSGLVATIAGGHLCSC CREERGTSGLVATIAGGHLCSC CREERGTSGLVATIAGGHLCSC CREERGTSGLVATIAGGHLCSC CREERGTSGLVATIAGGHLCSC CREERGTSGLCTARRADGT LRETTLYQDNIQWFTGLAVDYNBRLTTAMADGT LRETTLYQDNIQWFTGLAVDYNBRLTLAMADGT LRETTLYQDNIQWFTGLAVDYNBRLTLISGMDEPH AIVVPLRGTMYWSDWGNPKETAMADGT LRETTLYQDNIQWFTGLAVDYNBRRTLISGMDCPH AIVVPLRGTMYWSDWGNPKETAMADGT LRETTLYQDNIQWFTGLAVDYNBRRTLISGMCCWL CLLSPSGPVCTCPNGRLDGGCCAASPSG MPTCRCTGTGTGFCCCTAASPGG MPTCRCTGTGTGFCCTAAGGGCCTAASPGG MPTCRCTGTGTGFCCTAAGGGCCTAASPGG MPTCRCTGTGTGFCCTAAGGGCCTAASPGG MPTCRCTGTGTGFCCAAGGGCCTAAGGGCCAASPG MPTCRCTGTGTGFCCAAGGCCAASPG MPTCRCTGTGTGFCCAAGGCCAASPG MPTCRCTGTGTGFCCAAGGCCAASPG MPTCRCTGTGTGFCCAAGGCCAAGGCC MPTCCCAAGGGCCTAAGGCCAASPG MPTCRCTGTGTGFCCAAG			1				
GODDCGDSSDERGCHINECLSRKLSGCSQDC EDLKIGFKCRCRFGFLKDRCTADVDECS TTFPCSQRCINTHGSYKCLCVEGYAPRGGDP HSCKAVTDEEPFLIPANRYVIRKLINDGSNYY TLLKQGLNNAVALDFDYREQMITWTDVTTQ GSMIRRMHLNGSNYQULHRTGLSNPDGLAV DWVGGNLYWCDKGRDTTEVSKLNGAYRTVL VSSGRERALVVDVQNGYLYWTDWGDSL IGRIGMDGSSSSVJVDTKTIWFRGLTLDYVTE RIYWADAREDYIEFASLDGSNEHVLSQDIPH IFALTLEEDYVYMTDWETKSINRAHKTTGTN KTLLISTLHRPMDLHVFHALRQPDVPHIPCK VNNGGCSNLCLISPGGGHKCACTFNYLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGHSDEPPDCEFKCRPGGPQCSTGICTN PAFICDGDNDCQDNSDEANCDHVCLPSQFK CTNTNRCPGFRCNGQDNCGDEBEDCPC TTCAPNGPOCESTIKRCDRVVVCDRDNDCVD GSDEPANCTQMTGVDEFRCXDSGRCPARW KCDGEDDCGDGSDEPKECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCCDDNSDESSCTPR PCSESEFSCANGRCIAGRWKCDGBEDCADGS DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMGSSBEAAGTGVRTCPLDEFPCANNT LCKPLAWKCDGEDDCGDNSDESTPR CPSESEFSCANGRCIAGRWKCDGBEDCADG GTDEEDCEPPTAHTTHCKDKKEFJCRNNRCL SSLKCNMFDDCGDSDSDENECARFV CPPNFFRCKNBRVCLWIRGCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKFFJCRNNRCL SSLKCNMFDDCGDSDSDEDSUPSCANC USSLCRNGTDGGDGSDEDSUPSKINSCAT NASIGGBARCVRTEKAAYCACRSGFHTVPQ QPGCQDDMECLRETGTSQLCNTGGHLCSC ARNPMKTHNTCKAEGSEYQVLYIADDNERS LFPGHHSAYEQAFQODESVIRDAMDVIVKA GRYVTMWHTGTISTRSLIPAAPPTTSNRHR RQIDRGVTHLNISGLKERPKLAMGNY WTDSGRDVEWAQMKGGPNSKTLISGMDEPH AIVVPJLRGTMYWSDWGNIPKETAMDGT LRETTLYQDNIQWPTGLAVDYNNRHRETAMDGT LRETTLYQDNIQWPTGLAVDYNNRHRETAMDGT LRETTLYQDNIQWPTGLAVDYNNRHRETAMDGT LRETTLYQDNIQWPTGLAVDYNNRHRETAMDGT LRETTLYQDNIQWPTGLAVDYNNRHRETAMDGT LRETTLYQDNIQWPTGLAVDYNNRKYCHMFSLVVLTGG LSHASDVVLYHQHKQPEVTNPCDRKCCWL CLLSPGGPCCTPNGRLDNGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPCCLPGGGCTLNARRQPCCQ PRYTGDKCELDQCWERGRGGCTAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPCCLPGGTGGGCTLNARRGPCCQ PRYTGDKCELDQCWERGRGGCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPCCLPGGTGGGCTLNARRGVCRA CLTCVGHCSNGGGCTLNARRAMGPCQCPPIM TGRCEEHVFSQQQFGHLASILIPLLLLLLLLLLLV VAGVYFWYKRRVQQAKGGGPKTNAGAMPT	1	}	l				NPHCTSPEHKCNASSQFLCSSGRCVAEALLCN
TTFPCSQRCINTHGSYECLCVEGYARRGGDP HSCKAVTDEEPFLIFANFURLNILDGSNY TLLKQGLNNAVALDFDYREQMIYWTDVTTQ GSMTRRMHLNGSNVQVLHRTGLSNFDGLAV DWVGGNLYWCDKGRDTEVSKLNGAYRTVL VSSGLREPRALVVDQNGYLYWTDWGDHSL IGRIGMGDSSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVYLSQDIPH IFALTLFEDYYWTDWETKSINRAHKTTGTN KTLLISTLHRRMDLHVFHALRQPDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD GRTCVSNCTASQPVCKNNKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGQPCCSTGICTN PAFICOGIONDCQDNSDEACHOUNTCLSPGK CTNTNRCIPGIFRCNGQDNGCGDGEDCDCPE VTCAPNOPCCSTTRCIPRYWVCDRDNDCVD GSDEPANCTOMTGQVDEFRCKDSGRCIPARW KCDGEDDCGDSSDEPKECDERTCEPYOFRC KNNRCVPGRWQCDYDNDCGDNSDESSCTPR PCSESFSCANGCLAGRWKCDGDHDCADGS DEKDCTPRCDMDGPCCKSGRCIPARW KCDGEDDCGDSSDEPKECDERTCEPYOFRC KNNRCVPGRWQCDYDNDCGDNSDESSCTPR PCSESFSCANGCLAGRWKCDGDHDCADGS DEKDCTPRCDMDGPCCKSGRCIPLRWRCDA DADCMDGSDEACCTGVRTCPLDEPCCNTT LCKPLAWKCDGEDDCGDNSDEPECARPV CPPNRPFRCKNDRVCLWIGNQCDGTDNGGD GTDEEDCEPTAHTTHCKKFFLCNGNCL SSSLRCNMPDDCGGGSBDEDCSDPFECNRC CPPNRPFRCKNDRVCLWIGNQCDGTDNGGD GTDEEDCEPTAHTHTINGKFFLCNGNCL SSSLRCNMPDDCGGGSBEDCSDPFLTSCAT NASICGDEACRVTEKAACCRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLADDNERS LFFGHPISAFCQAFGGBSVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRIRR RQIDRGVTHINSGLKMPAKTGADWAGNYY WTDSGRDVEVAQMKGENRTLLIADDNERS LFFGHPISAFCQAFGGADWAGNYY WTDSGRDVEVAQMKGENRTLLIADDNERS LFFGHPISAFCQAFGGADWAGNYY WTDSGRDVEVAQMKGENRTLLIAGMDEPH ALVVDPLRGTMYMSGLMCRGADWAGNYY WTDSGRDVEVAQMKGENRTLLIAGMDEPH ALVVDPLRGTMYMSGLAFUNTCKGMDEPH ALVVDPLRGTMYMSGLAFUNTCKGMDEPH ALVVDPLRGTMYMSGLAFUNTCKGMDEPH ALSVOSRILNGTDFIVADGRACDWAGNYY WTDSGRDVEVAQMKGENRTLLIAGMDEPH ALSVOSRILNGTDFIVADGRACDWAGNYY WTDSGRDVEVAQMKGENRTLLIAGMDEPH ALSVOSRILNGTCPYPSSTPDV PEDYTYGYTYNNRVFKHKFGHSRLVALDA KLSVGSRILNGTDFIVADGRACDWAGNYY WTDSGRDVEVAQMKGENRTLLIAGMDEPH ALSVOSRILNGTCPYPSSTPDV PEDYTYGYTYNNRVFKHKFGHSRLVALDA KLSVGSRILNGTDFIVADGRACDWAGNYACCGYCE OPRYTGBKCELDQCWERGGTAWAGNGTANSTCT VNQGNOPCCCLOPGTGARGDTANSTCT VNQGNOPCCCLOPGTGARGTCAASSGG MFTCCCFTGGTGFKCTQQCAGYCANSTCT VNQGNOPCCCLOPGTGARGCTCAASSGG	ļ						GQDDCGDSSDERGCHINECLSRKLSGCSQDC
HSCKAVTDEPFLIFANRYYLRELNLDGSNY TILKQCILNWAVALDFDYEQMYWTDVTTQ GSMRRMHLNGSNVQVLHRTGLSNPDGLAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSSGLREPRALVVDVQNGVLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALTLFEDYVYWTDWETKSINRAHKTGTIN KTLLISTLHRPMDLHYPHALRQPDYPNHPCK VNNGGCSNLCLLSPGGHKCACPTNYYLGSD GRTCVSNCTASGTVCKNDKCIPFWKCDTE DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN PAFICDGDNDCQDNSDEANCDHYCLPSQFK CTNTNRCIPGFRCNGQDNCGDGEDERDCPE VTCAPNGPQCSITKRCIPRVWVCDDRDDCVD GSDEPANCTOMTGGVDEFFRCKDSGRCIPARW KCDGEDDCGDGSDEPKECDERTCEPYQFRC KNNRCVFGRWCQTYDNDCGDNSDESCTTR PCSSESFSCANGRCIAGRWKCDGDHDCADGS DEKDCTFRCDMDQPQCKSGHCIPLRWCDA DADCMDGSDEAACTIGVTGVTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENEECARFV CPPNPFRFRCKNDRVCLWIGNQCDGTDNCOD GTDEEDCEPPTAHTTHCKDKKEFILRNCAL SSSLRCNMFDDCGDGSDEEDCSIDFLTSCAT NASICGDEACTGYTCFLUGRQCDGTDNCOD GTDEEDCEPPTAHTTHCKDKKEFILRNCAL SSSLRCNMFDDCGDGSDEEDCSIDFLTSCAT NASICGDEACVATEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNYMKTHNTCKAEGSSYQVLYIADDNERS LFFGHPBSAYEQAFQGDESVRIDAMDHVKA GRYYMTNWHTGTISTSRAYPAPPTTSNRFR RQDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVEVAQMKGENRTLLISGMIDEPH AIVVDPLRGTMYWSDGWGHPKIETAAMDGT LRETLVQDNQWFTGLAVDYHNERLYWADA KLSVIGSRIKJGTDFIVAGMCHFKIETAAMDGT LRETLVQDNQWFTGLAVDYHNERLYWADA KLSVIGSRIKJGTDFIVAGMCHFKIETAAMDGT LRETLVQDNQWFTGLAVDYHNERLYWADA KLSVIGSRIKJGTDFIVAGNGCHESTENDV FEDYIYGVTYNNRVFKHRFGHSPLVALTGG LSHASDVVLYHQHKQPEVTNPCDRKCCEWL CLLSPSGPVCTCPNGKRLDNGTCVPPPSPTPP PDAFRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCLDQCWEHRGGTCANSFGT WFTCCFTGFTGFKCTQQVCAGYCACSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGTGCSGGSCFCNNSKKMMPECQCPPHM TGPRCEEHVPSQQQPGHHASLILLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL	ł						EDLKIGFKCRCRPGFRLKDDGRTCADVDECS
TLLKQGLNNAVALDFDYREQMYWDTVTIQ GSMRRMHLMGSNYQULHRGLSNPDGLAV DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSGALREPRALVVDVQNGYLYWTDWGDHSL IGKIGMGSSSSNYUDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVYLSQDIPH IFALTLFEDYYWTDWETKSINRAHKTTGTN KTLLISTLHRRMDLHVFHALRQPDYNHPICK VNNGGCSILCLLSPGGGRICACPTNFYLGSD GRTCVSNCTASQPVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGFQCSTGICTN PAFICDGDNDCQDNSDEANCDHYCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEBRDCPE VTCAPNOFQCSITRKCIPFWWCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDSDEPKECDERTCEPYGFRC KNNRCVPGRWQCDYDNDCGDNSDERSGTPARW KCDGEDDCGDSDEPKECDERTCEPYGFRC KNNRCVPGRWQCDYDNDCGDNSDERSGTPA PCSESFSCANGRCLAGRKCGBHDCADGS DEKDCTPRCDMDQFQCKSGFICPLRWRCDA DADCMDGSDEACGTGVATCPLDEPCANT LCKPLAWKCDGEDDCGDNSDENBESCARV CPPNRFFRCKNDRVCLWIGRQCDGTDNGGD GTDEEDCEPTAHTTHCKKEFLCRNQRCL SSLRCNMFDDCGGGSDEDECGDNSDENBEECARV CPPNRFFRCKNDRVCLWIGRQCDGTDNGGD GTDEEDCEPTAHTTHCKKKEFLCRNQRCL SSLRCNMFDDCGGGSDEDCGDNSDENDECARRV QPGCQDINECLRFGTCSQLCNNTKGGFLCSC ARNPMKTHNTCKAEGSEYQVLYIADDNEIRS LFFGHPBAYEQAFQGDESVRDAMDVHVKA GRYYWTNWHTGTISYRSLPPAAPPTINNRIR RQIDRGYTHINISGLKMFRGJADDWVAGNYY WTDSGRDVIEVAQMKGENRTLLISGMDEPH AIVVDPLRGTMYWSDGWFRFEIETAMMDGT LRETLVQDNQWFTGLAVDYHNERLYWADA KLSVGSRLGNTMYBNGGHPKEIETAMMDGT LRETLVQDNQWFTGLAVDYHNERLYWADA KLSVGSRLGNTMYBNGGHPKEIETAMMDGT LRETLVQDNQWFTGLAVDYHNERLYWADA KLSVGSRLGNTMYBNGGHPKEIETAMMDGT LRETLVQDNQWFTGLAVDYHNERLYWADA KLSVGSRLGNTMYTNYKGMFEGALDWVAGRNYY WTDSGRDVIEVAQMKGENRTLLISGMDEPH AIVVDPLRGTMYWSDGWFRFEIETAMMDGT LRETLVQDNQWFTGLAVDYHNERLYWADA KLSVGSRLGNTDPIVAGRCGLSCFFSIDV FEDYIYGVTYNNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPPSSTPP PDARRPGTCNLCCFNGGSCFLNARRQPKCRC QPRYTGNKCELDQCWERKGGGSTCTAXFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTGRYAPS CLTCVGHCSNGGSCTMSKKMMPECQCPPIM TGRCCEHVFSQQQPGHHASLILLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL				'			
GSMIRRMELNGSNYQVLHRTGLSNPDGLAV DWYGGNLYWCDKGRDTEVSKLNGAYRTVL VSSGLREPRALVVDVQNGYLVWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPIGLTLDYVTE RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALTLFEDYVYWTDWETKSNRAHKTTGTN KTLLISTLHRPMDLHVALRQPDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTNFVLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGQPQCSTGICTN PAFICOGDNDCQDNSDEANCDHVCLPSQFK CTNTNRCFGIFCROGQDNCGDGEDERDCPE VTCAPNQPQCSITIKCCIPRVWVCDRDNDCVD GSDEPANCTQMTCGPTRCKDSGRCFBAW KCDGEDDC3GDSDEKEECDEPRCKDSGRCFBAW KCDGEDDC3GDSDEKEECDEPRCKDSGRCFBAW KCDGEDDC3GSDEFKEECDEFTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCLAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCHLRWRCDA DADCMDGSDEAACGTGWRCDGDNSDESCTPR PCSESEFSCANGRCLAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCHLRWRCDA DADCMDGSDEAACGTGWRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDEPECARPV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDSDEEDSEPECARPV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDSDEEDSCHTLTSCAT NASICGDEARCVATEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHTCKAGESEYQVLYLADDNERS LFPGHPHSAYSQAPQGDESVRIDAMDVHVKA GRYWTNWHTGTISTSLPPAAPPTTSNRIR RQDRGVTHLNISGLKMPRGJADWVAGNVY WTDSGRDVEVAQMKGENNKTLISGMIDEPH AIVVDPLAGTMYWSDWGNEPKETAAMDGT LRETLVQDNIQWPTGLAVDYHRRLYWADA KLSVIGSRLNGTDPTVADSKRGLSHFSIDV FEDYTYGVTYNNRVFKHKFGHSPLVNLTGG LSHASDVVLYHOHGCYPVPSPTPP PDARPGTCNLQCNGGSCELNARCPKCCK QPRYTGDKCELDQVENCENGGTCCAASPSG MPTCRCFTGFTGPRCTQQVCAGNSTCT VNQGNQPCCLLPGTLGRCQYCACNSTCT VNQGNQPCCCLPGTLGRCQCYCACNSTCT VNQGNQPCCCLPGTLGRCQCYCACNSTCT VNQGNQPCCCLPGTLGRCQCYCACNSTCT VNQGNQPCCCLPGTLGRCQCYCACNSTCT VNQGNQPCCCLPGTLGRCQCYCACNSTCT VNQGNQPCCCLPGTLGRCQCYCACNSTCT VNQGNQPCCCLPGTLGRCQCYCACNSTCT VNQGNQPCCCLPGTLGRCQCYCACNSTCT VNQGNQPCCCLPGTLGRCQCYCACNSTCT VNGGNQPCCCLPGTLGRCQCYCACNTSTCT VNGGNQPCCCLPGTLGRCQCYCACNTSTCT VNGGNQPCCCLPGTLGRCQCYTCCTTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVPSQQQPGHLASILPLLLLLLLLLL VAGVVFWYRRRVQCQAKGGPGCTTMGAMMGAM	Ì						
DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL VSSGLREPRALVATOVGYLYWTDWGDHSL IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE RIYWADAREDYIEFASLDGSNRIVVLSQDIFH IF ALTILEPDYYWTUDETKSSNRAHKTTGTN KTLLISTLIRPMDLHVFHALRQPDYNHPCK VNNGGSNLCLLSPGGHKCACPTNFYLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN PAFICDGDNCQDNSDEANCDHVCLPSQFK CTNTNRCPGIFRCNGQDNCGDGEDERDCPE VTCAPNQPCSITIKRCIPTWWCDDTDMOCVD GSDEFANCTQMTGQVDETRCKDSGRCIPARW KCDGEDDCGDSDEPRCEPTQFFR KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGHDCADGS DEKDCTPRCDMDGPCKSGHCIPLRWRCDA DADCMDGSDEEAGTGVTCPLDEFQCNNT LCKPLAWCDGEDDCONSDENEPEGARV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCMRFDDCGDGSDEDENEPEGARV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDIDNELREGTGCUCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNERS LFPGHPHSAYEQAPQODESVRIDAMDVHVKA GRVYWTNWHTGITSYSLPPAAPPTTSNRIR RQIDRGVTHLNISGLKMPRGIADMVAGNVY WTDSGRDVIEVAQMKGENKRTLISGMDEPH AIVVDPLRGTMYWSDWGNEPKEITAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPTVAADSKRGLSIPPSIDV FEDYYGVTYNNRVFKHIKFGHSELVNLTGG SHASDVVLYHQHGVCPVNPCDRKKCCWL CLLSPSGPVCTCPNGRKLDNGTCAASPSG MPTCRCTFTGFTGPKCTQVCAGYCANNSTCT VNQGNQPQCRCLPFGLGDRCQYPACSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSQVCTNCTCDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVPSQQPGHLASILLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLLL		1	l				
VSSGLREPRALVYDVQNGYLYWTDWGDHSL IGRIGMGDSSSSVIJUTIVPNGCILDYYTE RIYWADAREDYIEFASLDGSNRHVVLSQDIPH IFALTLFEDYYYWTDWETKSNRAHKTTGTN KTLLISTLIRPMDLHVPHALRQPDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD GRTCVSNCTASQFVCKNDKCIPFWKCDTE DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN PAFICDGDNDCQDNSDEANCDHVCLPSQFK CTNTNRCPGIFRCNGQDNCGDGEDERDCPE VTCAPNQPQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSSEJSFSCANGRCIAGRWKCDGBHDCADGS DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMDGSDEAACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDEESCTPR CPPNRPFRCKNDRVCLWIGRQCDTDNCGD GTDEEDCFPTAHTTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDOSDEEDCSIDPKLTSCAT NASICGDEARCVTEKAAYCACRSGFHTVPQ QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQUTADDNEIRS LFPGHPBAYSQAFQODESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRIR RQIDRGVTHLNISGLKMPRGIADWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMDEPH AIVVDPLRGTMYWSDWGNBEKETAAMDGT LRETLVONNQWPTGAVDYHNERLYWADA KLSVIGSRLNGTDPIVAADSKGGLSIPPSIDV FEDTYGVTYINNRYFKIHKFGHSLEYAWADA KLSVIGSRLNGTDPIVAADSKGGLSIPPSIDV FEDTYGVTYINNRYFKIHKFGHSLEYAWADA KLSVIGSRLNGTDPIVAADSKGCLSHPFSIDV FEDTYGVTYINNRYFKIHKFGHSLEYAWADA KLSVIGSRLNGTDPIVAADSKGCSTLNARQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCTTGFTGFTCTQUCAGYCANNSTCT VNQGNQPQCRCLPGFLGBRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFGSRCEVNK CSRCLEGGCVVNKSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEBUTSQQQPGHLASILPLLLLLLLLLLLV VAGVVWYKRRVQQAKGGPQRTNGTAM	į.		l ·				
GRIGMDGSSRSVIJDTKITWPNGLTLDYVTE RIYWADAREDYIER SLDGSNRIVVLSQDIPH IFALTLFEDYVYWTDWETKSINRAHKTIGTN KTLLISTLIRPMDLHVFHALRQPDVPNHPCK VNNGGGSNLCLLSPGGHKCACPTNFYUGSD GRTCVSNCTASQFVCKNIKCTPFWKCDTE DDCGDHSDEPPDCFEFCRFGQFQCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIFRCNGQDNCGGEDERDCPE VTCAPNQFQCSITKRCFRVWVCDRDNDCVD GSDEFANCTQMTCGYDEFRCKDGRCPARW KCDGEDDCGDGSDEFKECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCIAGRWCDDGHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWCDGEDDCGDNSDENFEECARV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEDCEPPTAHTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDGSDEEDCSIDFKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCGLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNERS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRGNVEYAMAGGENKTLISGMDEPH AIVVDPLRGTMYWSDWGNHPKETAAMDGT LRETLVQDNIQWPTGLAVYHNELYWADA KLSVIGSRLNGTDPTVAADSKRGLSHPFSIDV FEDYYGVTYNNRVFKHKFGHSPLVNLTGG USHASDVVLYHQHKCGWLC CLLSPSGPVCTCPNGKRLDNGTCVPYPSTPP PDAPPPGTCNLQCTPGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGTCAASPSG MPTCRCTTGTFGKCTQVACAGYCANNSTCT VNQGNQPQCRCLPFGLGRCCYTRCGSCCE NFGTCQMAADGSRQCRCTAYFGSRCEVIA CSRCLEGGCVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVPSQQQPGHLASILPLLLLLLLLLLLV VAGVVFWYRRRVQQAKGFQRVTNGTMA					1		l
RIYWADAREDYIEFASLOGSNRHVVLSQDIPH IFALTLFEDYVYWTDWETKSINRAHKTTGTN KTLLISTLHRPMDLHVFHALRQPDVPNHPCK VNNGGCSNLCLLSPGGGHKCACPTNYTLGSD GRTCVSNCTASGPVCKDMCCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGGPCCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITRCIPRVWVCDRDNDCVD GSDEPANCTQMTCOMPERCKDSGRCIPARW KCDGEDDCGDGSDEPKECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCLAGRWKCDGHDLADGS DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCNDRVCLWIGRCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVATEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQULYIADDNEIRS LFPGHPBSAYEQAFQGDESVRIDAMDVHVKA GRYYWTNWHTGTISYASLFPAAPPTTSNRHR RQDRGVTHLNSGLKMPRGIADWVAGNVY WTDSGRDVELVAGNKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGHNEELTAAMDCT LREILVODNIQWFGLAVDYHNERLYWADA KLSVIGSRLINGTDPPVAADSKRGLSHPFSIDV FEDYIYGVTYNNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHFKQFEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRFGTCNLCCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCANSTSCT VNGGNQPQCRCLAPFGSRCEVNK CSRCLEGACVVNKGGDVTCNTGGRASPSG MPTCRCPTGFTGPKCTQVCAGYCANNSTCT VNGGNQPQCRCLAPFGSRCEVNK CSRCLEGACVVNKGGSTTMNSKMMPECQCPPHM TGPRCEEHYFSQQQPGHIASLIPLLLLLLLVL VAGVVFWYKRRVQGAKGFORRMTNGAM							*
IFALTLEDYYYWTDWETKSINRAHKTTGTN KTLLISTLHRPMDLHVFHALRQPDVPNIFPCK VNNGGCSNLCLLSPGGHKCACPTNYTLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDEFKCRPQGPCGTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIFRCNQQDNCGDGEDERDCPE VTCAPNQFQCSTIKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGHDLCADGS DEMCTPRCDMDQCKSGHCIPLRWRCDA DADCMDGSDEAACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRFFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMPDDCGDGSDEDCSDPELTSCAT NASIGGDEARCVRTEKAAYCACRSGFHTVPQ QPGCQDINECRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEVQVLTADDMEIRS LFPGIPHSAYEQAFQGDESVRIDAMDVHVKA GRYYWTNWHTGTISYRSLPAAPPTTSNRHR RQDDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWDGMSPPKEITAAMDGT LREILVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYNNRYKHKFGHSFLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVSPPTPP PDAPRFGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGFCQCQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGGACVVNKYQQAKGFORKMMPECOCPPHM TGPRCEEHVFSQQQPGHLASILPILLLLLLVL VAGVVFWYKRRYQGAKGFQRRMTNIGAM					-	'	
KTLLISTLHRPMDLHVFHALRQPDVPNHFCK VNNGGCSNLCLLSPGGHKCACPTNY*LGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPDCPEFKCRPGQFQCSTGICTN PAFICDGDNDCQDNSDEANCDHIVCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGDSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDGFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRFPRCKNDRVCL WIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDSIDPKLTSCAT NASICGDEARCVATEKAAVCACRSGFHTVPG QPGCQDDNECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAPPTTSNRHR RQDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLAGTMYNSDWGNEPKETAAMDCT LREITLVQDNIQWPTGLAVDYNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGTYTINNTSYKHLKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRFGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGFKCTQQVCAGYCANNSTCT VNGGNQPQCRCLAFGLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKSGDVTCNCTTOGRVAPS CLTCVGHCSNGGSCTLNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHLASLIPLLLLLLVL VAGVVFWYKRRVQGAKGFORQRMTNGAM			1				
VNNGGCSNLCLLSPGGHRCACPTINFYLGSD GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEFKCRPGGFQCSTGICTN PAFICDGDNDCQDNSDEANCDHVCLPSQFK CTNTTNCCRGIFFCROQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKECDEPTCPFYQFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCLAGRWKCDGDHDCADGS DEKDCTFRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGDDCGDNSDEBSCTRP PCSESEFSCANGRCLWGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSDNFLISCAT NASICGDEARCVRTEKAAVCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNIPMKTINTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENKTLISGMDEPH AIVVDPLRGTMYWSDWGNFIPKETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSRINGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAAPRGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCCRTPGTPGKCTQQVCAGYCANISTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQVAGKGFQHQRNTNIGAM	1	1	i		ł		
GRTCVSNCTASQFVCKNDKCIPFWWKCDTE DDCGDHSDEPPDCPEKCRPGQFQCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLFSQFK CTNTNCCPGIFRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESSCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEFCCTPR CEPPREFRCKDDQFQCKSGHCIPLRWRCDA DADCMDGSDEGTGTRATTHCKDKKETLCRNQRCL GSSLRCNMFDDCGDGDSDEPECARFV CPPNRFFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKETLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSUPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTENTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRYVWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIADDWAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAUPYHNELYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCCCTFGTFGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHLASILIPLLLLLLLLV VAGVVFWYKRYQQAKGPQQRGTNGNAGM					į		
DDCGDHSDEPPDCPEFKCRFGQFQCSTGICTN PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIRCNGQDNCGDGEDERDCPE VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCFARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDESSCTPR PCSSEFSCANGRCIAGRWKCDGHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENFEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTERAAYCACRSGFHTVPG QPGCQDNECLRTEKAAYCACRSGFHTVPG QPGCQDNECLRTEKAAYCACRSGFHTVPG QPGCQDNECLRTEKAAYCACRSGFHTVPG QPGCQDNECLRTEKAAYCACRSGFHTVPG QPGCQDNECLRTETGAYCACRSGFHTVPG QPGCQDNECLRTETGAYCACRSGFHTVPG QPGCQDNECLRTETGAYCACRSGFHTVPG QPGCQDNECLRTETGAYCACRSGFHTVPG QPGCQDNECLRTETGAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG QPGCQDNECLRTETAAYCACRSGFHTVPG LSCACRGFTAAYCACRSGFHTVPG LSCACRGFTAAYCACRSGRACACRACRACRACRACRACRACRACRACRACRACRACRA					ĺ		
CTNTINCIPGIFECNGODNCGDGEDERDCPE VTCAPNOFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDCGDGSSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTKAGGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQQDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIADWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIFKIETAAMDGT LRETLVODNIQWFTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYNNRVFKHIKFGHSPLVNLTGG LSHASDVVLYHQHKQFEVTVPPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDARPRGTCNLQCFNGGSCTLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMSKMMPECQCPPHM TGPRCEEHVFSQQQPGHASILIPLLLLLLLLLL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							
VTCAPNOFQCSITKRCIPRVWVCDRDNDCVD GSDEPANCTQMTCGVDEFRCKDSGRCIPAR W KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCYPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCL WIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNORCL SSSLRCNMFDDCGDGSDEEDCSIDFKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAPQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIPKETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSRINGTDFIVAADSKRGLSHFFSIDV FEDYIYGVTYINNRVFKHKFGHSFLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MFTCRCFTGFTGFRCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLLL VAGVVFWYKRRVQGAKGFQHQRMTNGAM					ļ		PAFICDGDNDCQDNSDEANCDIHVCLPSQFK
GSDEPANCTQMTCGVDEFRCKDSGRCIPARW KCDGEDDCGDGSDEPKEECDERTCEPYQFRC KNNRCVPGRWQCDYDNDCGDNSDEESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKNCKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDDELRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKHIKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCCPNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGBRCQVRQCSGYCE NFGTCQMAADGSRQCRCCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHLASILIPLLLLLLLLL VAGVVFWKRRVQQAKGFQHLASILIPLLLLLLLLL VAGVVFWYKRRVQQAKGFQHLASILIPLLLLLLLLL VAGVVFWYKRRVQQAKGFQHLANSILIPLLLLLLLLL VAGVVFWYKRRVQQAKGFQHAMSILIPLLLLLLLLL VAGVVFWYKRRVQQAKGFQHQRMTNGAM	1	ľ	ł	ĺ	[
KCDGEDDCGDGSDEPKEECDERTCEPYOFRC KNNRCVPGRWQCDYDNDCGDNSDESCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRFPRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNERS LFPGHPHSAYEQAFQDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVUPPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHFPSIDV FEDYINGTYTNNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDARPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL	-						VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD
KNNRCVPGRWQCDYDNDCGDNSDESSCTPR PCSESEFSCANGRCIAGRWKCDGDHDCADGS DERDCTPRCDMDQFQCKSGHCPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNERS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRYYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLVL			l				
PCSESEFSCANGRCIAGRWKCDGBHDCADGS DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCL WIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCMMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNPMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYITYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGGSCFLNARQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGWVFWYKRRVQQAKGPQRRMTNGAM			1				-
DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA DADCMDGSDEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENPEECARFV CPPNRPFRCKNDRVCL WIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQODESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQQAKGPQRRMTNGAM	}	ļ	J	j			1
DADCMDGSDEACGTGVRTCPLDEFQCNNT LCKPLAWKCDGEDDCGDNSDENFEECARFV CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD GTDEEDCEPFTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDFKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNIEPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDARRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCTFGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1						
LCKPLAWKCDGEDDCGDNSDENPEEČARFV CPPNRPFRCKNDRVCL WIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGROVIELNAGMKGENKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKLHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM				i			
CPPNRPFRCKNDRVCL WIGRQCDGTDNCGD GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRFGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1						
GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVPSQQQPGHIASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	ĺ				l		
SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPPPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1	1	ł		l		·
NASICGDEARCVRTEKAAYCACRSGFHTVPG QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQPGHIASILIPLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	l l				İ		
QPGCQDINECLRFGTCSQLCNNTKGGHLCSC ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPITSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKLHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			1	1	ł !		
LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYTYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1			l			
GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM							
RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1			1			
WTDSGRDVIEVAQMKGENRKTLISGMIDEPH AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGTGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	Ì	1	1	1	ł		
AIVVDPLRGTMYWSDWGNHPKIETAAMDGT LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM]		
LRETLVQDNIQWPTGLAVDYHNERLYWADA KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	-]	ļ		
KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM			İ	1	1		
FEDYTYGVTYINNRVFKIHKFGHSPLVNLTGG LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM				[
LSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	1			ĺ	1		
CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM]			
PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM				1			
QPRYTGDKCELDQCWEHCRNGGTCAASPSG MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM						1	
MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM]		}	ļ	1		
VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM				1		1	
NFGTCQMAADGSRQCRCTAYFEGSRCEVNK CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM]	1		
CSRCLEGACVVNKQSGDVTCNCTDGRVAPS CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM				i	1		NFGTCQMAADGSRQCRCTAYFEGSRCEVNK
CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL VAGVVFWYKRRVQGAKGFQHQRMTNGAM	ľ		1	1	· ·		CSRCLEGACVVNKQSGDVTCNCTDGRVAPS
VAGVVFWYKRRVQGAKGFQHQRMTNGAM			1	1	ł		CLTCVGHCSNGGSCTMNSKMMPECQCPPHM
				1			
NVEIGNPTYKMYEGGEPDDVGGLLDADFAL			1		!		
				<u> </u>			NVEIGNPTYKMYEGGEPDDVGGLLDADFAL

NO: of nuclectide peptide sequence NO: of nuclectide sequence NO: of nuclectide peptide sequence NO: of nuclectide peptide sequence NO: of nuclectide peptide sequence NO: of nuclectide peptide sequence NO: of nuclectide peptide sequence NO: of nuclectide peptide sequence NO: of nuclectide peptide sequence NO: of nuclectide peptide sequence NO: of nuclectide location corresponding to last amino acid residue of peptide sequence NO: of nuclectide location corresponding to last amino acid residue of peptide sequence NO: of nuclectide location corresponding to last amino acid residue of peptide sequence NO: of nuclectide location corresponding to last amino acid residue of peptide sequence NO: of nuclectide location corresponding to last amino acid residue of peptide sequence NO: of nuclectide location corresponding to last amino acid residue of peptide sequence NO: of nuclectide location corresponding to last amino acid residue of peptide sequence yellow.	H=Histidine, eucine, e, P=Proline, =Serine,
eotide sequence USSN 09/496 correspondi to last amino acid residue of peptide USSN 1-Letter (Sequence) 1-L	eucine, e, P=Proline, =Serine,
sequence 09/496 correspondi to last amino M=Methionine, N=Asparagin quence 914 ng to first amino acid residue Q=Glutamine, R=Arginine, S=Threonine, V=Valine, W=7	e, P=Proline, =Serine,
uence 914 ng to first acid residue Q=Glutamine, R=Arginine, S= amino acid of peptide T=Threonine, V=Valine, W=T	=Serine,
amino acid of peptide T=Threonine, V=Valine, W=	
	l ryntonhan
peptide /=possible nucleotide deletion	
sequence nucleotide insertion	, · possiois
DPDKPTNFTNPVYATLYM	GGHGSRHSLASTD
EKRELLGRGPEDEIGDPLA	
374 1724 A 3187 191 1815 CLELASAGKIPEESKALSLI	
AGLLPIPTPNPLTTLGVSLS	
ATLGEIPQPPLMGNVDPSK	IDEIRRTVYVGNL
NSQTTTADQLLEFFKQVGI	
QPTRFAFVEFADQNSVPRA	
LKINHSNNAIVKPPEMTPQ	AAAKELEEVMKR
VREAQSFISAAIEPGWLHS'	
RMYRE*APCTICGTFHLCL	
K*FFPPRVWKEQ*KKRR\R	SRSHTRSKSRSSSK
SHSRRKRSQSKHRSRSHNF	RSRSRQKDRRRSK
SPHKKRSKSRERRKSRSRS	HSRDKRKDTREKI
KEKERVKEKDREKERERE	KEREKEKERGKN
KDRDKEREKDREKDKEKI	DREREREKEHEKD
RDKEKEKEQDKEKEREKD	RSKEIDEKRKKDK
KSRTPPRSYNASRRSRSSSI	
PRTSKTIKRKSSRSPSPRSR	
HISERRERESTSMRKSSNI	DKDGKEKLEKNSI
375 1725 A 3192 415 101 AHSSHQTRAILQEFQWDIII	DEPOSIT AT CC
FIFPNLKKSLRGTHFSSVK	ATTI TUI MEADD
WF/FFYP*SPDLQIPSSFRNO	
PDLDGAYVKK	JEIND W THISQRC
376 1726 A 3199 931 418 GV*WCDLGSPQPPPPGFKQ	FCI GRSSSWDYR
HVPPHPANFVFLLETGFLH	
ASQSAGITGVSHTWPKNH	
K	
377 1727 A 3201 274 1285 KTGYTSRGSPLSPQSSIDSE	LSTSELEDDSISM
GYKLQDLTDVQIMARLQE	ESLRQDYASTSAS
VSRHSSSVSLSSGKKGTCS	DQEYDQYSLEDEE
EFDHLPPPQPRLPRCSPFQF	
RRSPSSQYFPSNNYQQQQY	YSPQAQTPDQQP
NRTNGDK/PPKKYA*PSPD	
VTVRNSQSFDSSLHGAGNO	
QHRVHSVGHFPVSIRQPLK	
NMPLSNGLQLYSNTGIPTP	
ALPRPSLAINGSNLPRSKIA	QPVRSFLQPPKPL
	11 0000010
1.57.20.20.20.20.20.20.20.20.20.20.20.20.20.	
YKGFVHKVELDRVKLSFS PEVVALEV TENDODI DAOL	1
PFKVNFY/TFNRQPLRV\QF PMLFP\VAPRDVPLLPSDVI	MALELIGKWLLW
EQLQAMRHIVTGTTRPAPY LVEAIKQVVKHLPKAHILA	
QRLRVHLPSSIYRLLAPSRI	
WDAKKGEYVFPAKKKLQI	
LVSAQFPIDHFTHIFIDEAG	HCMEDES! VATAC
LMEVKETGDPGGQLVLAG	
TQKHGLGYSLLERLLTYNS	
FITKLLRNYRSHPTILDIPNO	
DVVDRERFCRWAGILPRQO	
EREGNSPSFFNPEEAATVT	SYLKLLLAPSSKK
GKARLSPRSVGVISPYRKQ	
ELRGLDDIKDLKVTCCSTV	
ETSSSFHSSPRPRPTPAALN	RARALPEPLTPGD
SNLRVWDGIRKPACLTNTS	SCHS
379 1729 A 3206 432 130 PKAAPSVXLWFPPFL*GSFI	KPTKGHTXCVXIK
*LSTREAXDSXPGRQIAXX	RQGGKVETTTAL

No. of No. of No. of No. of Incidence of coation of the coation	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Couling Courseponding Part Courseponding Part Courseponding Part Courseponding Part Courseponding Part P		NO: of	hod				D=Aspartic Acid, E=Glutamic Acid,
1730			ľ				
Service Serv		•		4		to lest amino	
amino acid residue of peptide residue of peptide peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide insertion	-	401100		1			
residue of peptide pep]				
peptide	1		1				Y=Tyrosine, X=Unknown, *=Stop codon,
	i		1		peptide	•	
380					sequence		
1730			-				
PEGRIPHAPHKQRAGQGLFÖPQAWGGFKQAXLLERAQASSGVTPNG							
AXILIPGINGGRÉPVASLEARAQÁSSGVTPNG GGRTPYPTPSSGE	380	1730	A	3207	187	507	
381 1731 A 3225 1 840 GGRTYPYPTFSGE EMQLITSLGLQEFDIARNULELIYAQTILVMIQ EMQLITSLGLQEFDIARNULELIYAQTILVMIQ FEPCILIFRIQMIA FIMPTSKINISLEMMPGPPS KAWRASQMMTFFFILIFPSFTGVLCTLATIT WRIKEPSADCGFFGULFHISHYSWIDTLSTRP GYLWVVWIVRILGSVHFFFILITLVILITYLY WQTTEGRKIMIRLIHEQIINECKDKMFLJEKLI KLQDMEKKANPSSLYLERGEVGQGFLHLGE HDOSLDLRSRRSVQGGNPRA 382 1732 A 3238 256 38 LLMIKVSSTGTSGLHHIHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH]						
1731 A 3225 1 840 GTRPGHLPAPSDGFCVHL*SIPSWGSFGGSLJ							
EMQLITS.GL.QEFDLARNVLELITYAGTI.VWIGI FFCPLLFPTQMMEJIMFSNSMISLAMMFQPPS KAWRASQMMTFFILLFFSFTGVLCI.ATI WRLKPSADCOPPRGI.PJFISHSYSWIDTLSTRP GYLWVWYRNLIGSVHFFFILTI.IV.IITYLY WQITEGRKIMIRLIHEQIINEGKDKMFLIEKLI LQDMEKKANPSSI.VLERREVEQQGFLIHLIG HDGSLDLARRRSVQEGMPRA 382 1732 A 3238 256 38 LLMIKVSSTCFSCHLEHHHHHHHHHHRHHQGHNS LFFSLKSSSNSSTLPVVLSYNIILVFSKCLVPDF LFSNACL SPSNACL SKAMPASVALVALVSTADTUNKTGASSYG EQTI.HYLATNGESAVVQLPKNGPHYDVVWNS SSTEPCAVVGPMAAATIPLTOSDHKTGASSYG EQTI.HYLATNGESAVVQLPKNGPHYDVVWNS SSTEPCAVVGPMAAATIPLTYSSLHKXDVPS NAELWQVSWQPFLDGIPAKTITTQAVPSEVP NEEPKVATAYPPALRNKITYTGSLHKXDVPS NAELWQVSWQPFLDGIPAKTITTYQAVPSEVP NEEPKVATAYPPALRNKITYTSSLHKXDVPS NAELWQVSWQPFLDGIPAKTITTYQAVPSEVP NEEPKVATAYPPALRNKITYTSSLHKXDVPS NAELWQVSWQPFLDGIPAKTITTYQAVPSEVP NEEPKVATAYPPALRNKITYTGSLHKXDVPS NAELWQVSWQPFLDGIPAKTITTYQAVPSEVP NEEPKVATAYRPALRIKRITYQSLISGDP EIDKKINALKKLKAEQLKEQAATGKQLEK NQLEKIQKETALLQELEDELI TATCABRLRVNGVSKRELQELFTCTLFVIAALAAVFVDSAIP RLTTQHRPQDGSPFYTLLDPPLYQCCABPQQ LSQCARRVHGEKLRPTFGPRHRGAGTAKMS ASLVRATVRAVSKRRLQPTRAALTLITSANV KIKQLLKDKPEHYGVKYGVNRGCNGLSYTL EYYKTKGDSDEEVIQDQVRVFIEKKAQLTLL GTEMDYYEDKLSSEFVNNDKTGCTGGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPQLIFLILLSCFWIQGLPL KEEELPPGSETPTVASCALAELLHGALLER GPGMGYLPPGPGGGGEGTHTTTTTTTTTTTTTTTTTTTTTTTTTTTT	381	1731	A -	3225	1	840	GTRECHT PAPSINGECV/HI *SIPSWGSF*GESI /
FFCPLLPFIQMMLFIMFYSKNISLMMYRQPS KAWRASQMMTFFFILLIFSFIGVLCTLAITI WRLKPSADCGPPRGLP,FHSIYSWIDTLSTRP GYLWVVWYYRNLGSVHFFFILTLIVLITTYLY WQITFGRIKMRLLHEQINEGKDKMFJEKTLI KLQDMEKKANPSSLVLERREVEQQGFLHGE HDGGLDLSRRSVQGOPPRG. 382 1732 A 3238 256 38 LLIMIKVSSTCFSCHLHHHHHHHHRHHHHGHNS LFFSKSSNSSTLPVVLSVNILVFSKCLVPDF LFSNACL KGAPSFVRLYQYPNFAGPHAALANKSFFKAD KYTMLWNKKATAVLVIASTDVDKTGASYYG EQTLHYIATNOESAVQLFKNGPIYDVWNN SSTEPCAVYGFMAKATIFNLKCDPVFDFGTG PRNAAYYSPHGHILVLAGFGNLLIQI*ADIMK VWNKNYKLISKPVASDSTYVAVCPDGEHIL TATCAPRLRVNNGYKIWHYTGSILHKYDVPS NAELWQVSWQFFLOGIPFAKVCPDGEHIL TATCAPRLRVNNGYKIWHYTGSILHKYDVPS NAELWQVSWQFFLOGIPFAKVCPDGEHIL TATCAPRLRVNNGYKIWHYTGSILHKYDVPS NAELWQVSWQFFLSGTPAKNGKHEAKAAK QEARSDXSPDLAPTPAPQSTPRNTVSQSISGDP EIDKKIRNLKKKLKAIRQLKEQAATGKQLEK NQLEKIQKETALLQELEDLELGI 384 1734 A 3242 3 678 RSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSFYTILDPFLYLPGCQAPPQP LSCCARRVHGELARRPTFGPRHRGAGTAKMS ASLVATVRAVSKKLQTRAALTLTPSAVN KIKQLLKOKPPHVGVKVGVRTRGCNGLSVTL EYNTKTGGSDEGVIQDGVVTLYPLGGCAPPQP LSCCARRVHGELARPTFFGPRHRGAGTAKMS ASLVATVRAVSKKLQTRAALTLTPSAVN KIKQLLKOKPPHVGVKVGVRTRGCNGLSVTL EYNTKTGGSDEGVIQDGVVTLYPLGGCGGES FNI GTEMDYVEDKLSSEPVFNNPNIKGTCGCGES FNI GTEMDYVEDKLSSEPVFNNPNIKGTCGCGES FNI GTEMDYVEDKLSSEPVFNNPNIKGTCGCGES FNI GTEMDYVEDKLSSEPVFNGGGGGVLHFNCAGGGGTHAAT TLCRVSPEPGGAVGPNLTCKRWIEAAGGRTL VLAGGGSPGLAPLLANSSMLGEGQVLSPJF NRLLLHFQSPRVPRGGGGFRHYYGAYLLSCGFP PRRAHDDVSVTDLHPGGTATFHCDSGYQLQ EETLICLNGTRFSWNGETPSCMASCGGTHNA TLCRVSPEPGGAVGPNLTCKWVIEAAGGRTL HLHFERVSLDEDNDRLMVRSGGSPSPNYDS DMDDVPERGISDAQSLYVELLSETPANPLLL SLEFFAFEDRCFAPFLAHGNVTTTDPEYRPG ALATTSCLFQVA LEPPGPPNGLGQGFVLHFKEVPR NDTCPELPPPEGWGRTASHDAICVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPQODCVWGVHVQEEKRILLQVELINVREG DMLTLFFDGDGFSRRVLAQLKGHLRCVPPTEPHW NDTCPELPPPEGWGRTASHDGLIGGTVLTVQ EPDLTLLIPGGRGPGGGFVLHFKEVPR NDTCPELPPEGWGWTASHDGLIGGTVLTVC	301	1731		3223	•	040	
KAWRASQMMTFFILLFFFSFIGVLCLATI WRLKPSADGOPRGI P.FIFISFIGVLCLATI WRLKPSADGOPRGI P.FIFISFISWIDTLSTRP GYLWVVWYRNLIGSVHFFFIL TIJVLITTYJ Y WQITEGRKMIRULHEQINEGKOKMFLIEKLI KLQDMEKKANPSSI.VLERREVEQQGFLHLIGE KLQDMEKKANPSSI.VLERREVEQQGFLHLIGE HDGSLDLSRRSVQEGNPRA 382			1	<u> </u>	İ		
GYLWVWIYRRLIGSVHFFFILTLUITTYLY WOJTEGREKMRILHEQINSCKDEMFLIEKLI KLQDMEKKANPSSLVLERREVEQQGFLHLGE HDGSLDLRSRRSVQEGNPRA 1732 A 3238 256 38 LLMIKVSSTCSCLLHHIHHHHHHRHHQGHNS LFSLKSSSNSSTLPVYLSYNIILVFSKCLVPDF LFSNACL JFSNACL 383 1733 A 3241 1542 343 KGAPSFVRLYQYPNAGPHAALANSFFKAD KVTMLWNKKATAVLVIASTDVDKTGASYYG EQTLHYIATNGESAVVQLPKNGPIYDVWNS SSTEFCAVYGFMPAKATHFNLKCDPVPDFGTG PRNAAYYSPHGHLVLAGDILIQ("ADIMK VWNVKNYKLISKPVASDSTYFAWCPDGEHIL TATCAPRLRVNOFYKIMTYSSLHKVPVS NAELWQVSWQPFLDGIFPAKTTIYQAVPSEV NEEFKVATAYRPPALRIFNSKLHHEEPPP NMKPQSGNDKPLSKTALKNORKHEAKKAAK QEARSDKSPDLAPTPAPQSTPRNTYSQSISGDP EIDKKIKNLKKLKALGLEDLELGI 384 1734 A 3242 3 678 IRSPAARSPGLEFTCLLEVIAAIAAVFVDSAIP RLTOHRPQDGSPPYTILDPPLYLPGCAPPQR LSQCARRVHGEKLRRFTGPPHRGAGTAKMS ASLVRATVRAVSKKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVTRIGGCAPPQR LSQCARRVHGEKLRRFTGPPHRGAGTAKMS ASLVRATVRAVSKKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVTRIGGCABPQR FINI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPQLIFLILLSCPWIGGLP, KEEEILPFGSETPTVASEALAELHGALLRG GPEMGYLPGPPLGPEGGEETTTTITITTTTT TVTSPVLCNNNISEGGYVESPDLGSPVSRTL GLLDCTYSHTVYPGYGEIGEPTCLLTLSCPWIGGLP, KEEEILPFGSFTPTVASEALAELHGALLRG GPEMGYLPGPPLGPEGGEETTTTITITTTTT TVTSPVLCNNNISEGGYVESPDLGSPVSRTL GLLDCTYSHTVYPGYGEIGQVQTLNLSCGEELL VLAGGGSPGLAPRLLANSSMLGEGGVLRSPT NRILLHFGSPRVPRGGFTSPCMASCGGTHNA TLGRIVSPEPGGAVGPNLTICRWVIEAAEGRIL HLHFERVSLDENDRILMVRSGGSPSPVITDS DMDDVPERGLISDAGSLYVELJSSTPANPULL SLRFEAFEEDRCFAPFLAHONVTTTDPEYRPG ALATTSCLEVQYALEPGPPPRAILECVDPTEPHW NDTEPACKAMCGGELSSPAGVVLSSPDWPGS YSPGQDCVWGVPVQEEKRILLOVPELINVREG DMLTLFDGGPSAVLLAQLRGPCPPRRILLSS GPDLTLQFQAPPGPPRPRLLSS GPDLTLQFQAPPGPPRPRLLSS GPDLTLQFQAPPGPPRPRLAGGRILHREVPR NDTCPELPPPEWGWRTASGGDLIRGTVLTYQ							,
WOITEGREMMIRLIHEQIINSCKDEMPLIERLI KLQDMEKKANPSSLVLERREVFQQFFLHGE HDGSLDLRSRRSVQEONPRA			}				WRLKPSADCGPFRGLPLFIHSIYSWIDTLSTRP
RLQDMEKKANPSSLVLERREVEQQGFLHLGE HDGSLDLRSRRSVQEGORPRA							GYLWVVWIYRNLIGSVHFFFILTLIVLIITYLY
HIDÓSLDLRSRRSVQEGNPRA							
382							
LEFSLKSSSNSSTLPVYLSYNIILVFSKCLVFDF LFSNACL 1733 A 3241 1542 343 KGAPSPVRLYQYPNFAGPHAALANKSFFKAD KVYMLWNKKATAVLVIASTDVDKTGASYYG EQTLHYJIATNGESAVYQLPKNGPIYDVVWNS SSTEFCAVYGFMPAKATIFNLKCDPVFDFGTG PRNAAYYSPHGHILVLAGFGONLILQI*AD/IMK VWNVKNYKLISKPVASDSTYFAWCPDGHIL TATCAPRLRVNNGYKIWFIYTGSILHKYDVPS NAELWQVSWQPFLOGIFPAKTITYQAVPSEVP NEEPKVATAYRPALRNEPITNSKLHEEEPPO NMKPQSGNKPLSKTALQGIFPAKTITYQAVPSEVP NEEPKVATAYRPALRNEPITNSKLHEEEPPO NMKPQSGNKPLSKLAKQAKKAAK QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP EIDKKKNIKKAKALAGLKEPQAATGKQLEK NQLEKIQKETALLQELEDLELGI 384 1734 A 3242 3 678 IRSPAARSPGLETTPTCLLFVIAAIAVFVDSASI KIKQLLKDKPEHYGVKYGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPIKGTCGCGAPPOP LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS ASLVRATVRAVSKKLQPTRAAATTLTPSAVN KIKQLLKDKPEHYGVKYGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPIKGTCGCGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLIFLILLSCPWIQGLPL KEEELPERGSTETVASEALAELLHGALLRR GPEMGYLTGPPPGGEGEETTTTIITTTTTT TYTSPVLCNNISSEGEGGFRHTYGAYLLSCGEPP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGGGFRHTYGAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGGGFRHTYGAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGGGFRHTYGAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGGGFRHTYGAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGGGFRHTYGAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGGGFRHTYGAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGGGFRHTYGAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGGGFRHTYGAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGGGFRHTYGAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGGGFRHTYGAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGGGFRHTYTTPTPHY NDTCREACAMCGGELSEPAQCVLSPDWPQS YSPGQDCVGWGWYHQEEGGFLGTCHVITYGG DMLTLFDGGGPSARVLAQLRGPQPRRRLSS GPDLTLQFQAPPGPPPPGLGQGCVLHFKEVPR NDTCREPTPPEWGWKATASHGDLIRGTVLTYQ	222	1500	<u> </u>	2222		3.0	
LFSNACL 383 1733 A 3241 1542 343 KGAPSFVRLYQYPNFAGPHAALANKSFFKAD KVTMLWNKKATAVLVIASTDVDKTGASYYG EQTLHYIATNGESAVVQLFKNGPIYDVVWNS STEFCAVYGFNPAKATIFNLKCDPYDFDGTG PRNAAYYSPHGHLVLAGFGNLLIQI*ADIMK VWNVKNYKLISKPVASDSTYFAWCPDGEHIL TATCAFRLRVNNGYKIWHYTGSILHKYDVPS NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP NEPFKVATAYRPAKRYFINSKLHEEPPQ NMKPQSGNDKPLSKTALKNQRKHEAKKAAK QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP EIDKKIKNLKKLKAEQLKEQAATGKQLEK NQLEKQAATGKQLEKQAATGKQLEK NQLEKQAATGKQLEK NQLEKQAATGKQLEKGAATGKAA KAAX QEARDAYAGAATGAATGAATGAATAATAATAATAATAATAATAATAA	382	1732	A	3238	256	38	
383 1733 A 3241 1542 343 KGAPSPVRLYQYPNFAGPHAALANKSFFKAD KVTMLWNKKATAVLVIASTDVDKTGASYYG EQTLHYIATNGESAVVQLPKNGPIYDVVWNS SSTEFCAVYGPMPAKATIFNI.KCDPVPDFGTG PRNAAYSPHGHILV.AGFGNILD.QIF-ADIMK VWNVKNYKLISKPVASDSTYFAWCPDGEHIL TATCAPRILVNNGYKIWHYTGSILHKYDVPS NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP NEEPKVATAYRPPALRNKPITNSKLHEEPPQ NMKPQSGNDKPLSKTALKNQRKHEAKKAAK QEARSDKSPDLAPTPAPQSTPRNITVSQSISGDP EIDKKIKNI.KKKLKAIEQLKEQAATGKQLEK NQLEKIQKETALLGELDLELGI GI RISPAARSPGLETPTCLLFVIAAIAAVFVDSAIP RLIQHEPQDGSPPYILLDPPLYLPGQCAPPQP LSQCARVHGEKLENFTFOPRIRGAGTAKMS ASLVRATVRAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL BYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI VAMGTPRAQHPPPQLLFLILLSCPWIQGLPL KEEELLEPGSETPTVASSALAELLHGGALLRG GPEMGYLPGPPL.GPEGIGQVTLSPTT NRILLHFQSPRVPRGGGFRHIVQAYLLSGGPP RPPAHGDVSVTDLHPGGTATFHCDSGVQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRIL HLHFERVSLDEDNDRLMVRSGGSPLSPVITDS DMDDVPRRGLISDAQSLVVBLSETPANPLLL SLRFEAFEBERG-APFLAHGNVTTTDPEYNG ALATTSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVLSPUNPQS YSSGDCVWGVHQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRALLSS GPDLTLQFQAPPPPPNGLGGGFVLINFREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPPPPPGLGGGFVLIFKEVPR NDTCPELPPEWGWRTASHGDLIRGTVVLTYQ			1	!			j i
KVTMLWNKKĀTAVLVIASTDVDKTGASYYG EQTLHYIATNGESAVVQLPKNGPIYDVVWNS SSTEFCAVYGPMPAKATIFNLKCDPVFDFGTG PRNAAYYSPHGHILVLAGFGNLLQI*ADIMK VWNVKNYKLISKVSDSTYFAWCPDGEHIL TATCAPRLRVNNGYKIWHYTGSILHKYDVPS NAELWQVSWQPFLDGIPPAKTITYQAVPSEVP NEEPKVATAYRPPALRNKPITNSKLHEEEPPQ NMKPQSGNDKPLSKTALKNORKHEAKKAAK QEARSDKSPDLAPPALRNKPITNSKLHEEEPPQ NMKPQSGNDKPLSKTALKNORKHEAKKAAK QEARSDKSPDLAPPALRNKPITNSKLHEEEPPQ NMKPQSGNDKPLSKTALKNORKHEAKKAAK QEARSDKSPDLAPPALOSTPRATVSQSISGDP EIDKKIKNLKKKLKAIEQLKEQAATGKQLEK NQLEKIQKETALLQELEDLELGI 384 1734 A 3242 3 678 IRSPAARSPGLETPLLEVIAAIAAVFVDSAIP RLTQHPPQDGSPPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLTEVTAGATAWN KIKQLLKDKPEHVGVGVVRTRGCNGLSYTL EYTKTKGDSDEEVIDOPRVFTEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTITITTTTT TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSHVYPGYGEIGUVTLNLSQUEELL VLAGGGSPGLARPLLANSSMLGEGVVLSSTT NRILLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHADVSVTDLHGGTATFHCDSGVQLQG EETLICLNGTRPSWNGETPSCMASCGGTHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HHFERVISLDENDRILMVRSGGSSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAFFLAHGNVTTTDPEYPRO ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVYLSPDWPQS YSSGQDCVWGVHQEEKRILQVGLINVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPPNGLGQGFVLHYKEVPR NDTCPELPPPEWGWGRASHGDLIRGTVVLTYQ	383	1733	<u> </u>	3241	1542	3/13	
EQTLHYIATINGESAVVQLPKNOPIYDVVWNS SSTEPCAVYGFMPAKATIFNLKCDPVFDGTG PRNAAYYSPHOHILVLAGFGNLLQ *ADJMK VWNVKNYKLISKPVASDSTYFAWCPDGEHIL TATCAPRLRVNNGYKIWHYTGSLHKYDVPS NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP NEEPKVATAYRPPALRNIFNTSKLHEEPPQ NMKPQSGNDKPLSKTALKNQRKHEAKKAAK QEARSDK SPDLAPTPAPQSTPRNTVSQSISGDP EIDKKIKNLKKLKAIEQLKEQAATGKQLEK NQLEKIQKETALLQELEDLELGI 384	363	1755	Α	3241	1342	343	
SSTEFCAVYGFMPAKATIFNI.KCDPVFDFGTG PRNAAYYSPHGHIL VLAGFGONLIL.QI*AD/IMK VWNVKNYKLISKPVASDSTYFAWCPDGEHIL TATCAPRIRVNNGYKIWHTGSILHKYDVPS NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP NEEPKVATAYRPPALRNRYITINSKLHEEEPPQ NMKPQSGRDKPJLSTPAPQSTPRNTVSQSISGDP EIDKKIKNI.KKKLKAIEQLKEQAATGKQLEK NQLEKIQKETALLQELEDLELGI 384 1734 A 3242 3 678 IRSPAARSPGIETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLRPTFGPRHRGAGTAKMS ASLVATVRAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETFTVASEALAELLHGALLRR GPEMGYLFGPFGGGEETTTTITTTTTT TVTSPYLCNNNISEGEGGVVESPDLGSPVSRTL GLLDCTYSHVYPGYGEIQVQTILNLSQEEELL VLAGGGSFGLAPRLLANSSMLGEGQVLRSTT NRLLHFQSFRVGGGFRHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRFSWNGETPSCMASCGGTHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLVVELLSETPANPLLL SLRFFAFEEDRGCAPFLAHONYTTDPEYRFG ALATFSCLPGYALEPPOPNAIECVDPTEPHW NDTEPACKAMCGGELSPAGVVLSPDWPQS YSFGQDCVWGVHVQEEKRILLQVELINVREG DMLTLFDGGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ			1				
VWNVKNYKLISKPVASDSTYFAWCPDGEHIL TATCAPRLRVNNGYKIWHYTGSILHKYDVPS NAELWQVSWOPFLDGIFPAKTITYQAVPSEVP NEFPKVATAYRPPALRNKPITNSKLHEEEPPQ NMKPQSGNDKPLSKTALKNQRKHEAKKAAK QEARSDKSPDLAPTTAPQSTPRNTVSQSISGDP EIDKKIKNLKKKLKAIEQLKEQAATGKQLEK NQLEKIQKETALLQELEDLELGI RSPAARSFGLETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLRPTFGPPHRGAGTAKMS ASLVRATVRAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTIL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI VAMGTPRAQHPPPQLLFLILLSCPWIQGLPL KEEGILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTT TVTSPVLCNNISEGGGYVESPDLGSPVSRTL GLLDCTYSHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLILHFGSPRVPRGGGFRHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIENA TLGRIVSPFEGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRFG ALATTSCLPGYALEPPGPPNAIECVDPTEPHW NDTIEPACKAMCGGELSPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLLPGQAPPGPPPRGLQGGFVLHFKEVPR NDTICPELPPPEWGWRTASHGDLIRGTVVLTYQ	1						
TATCAPRLRVNNGYKIWHYTGSILHKYDVPS NAELWQVSWQPFLDGIPAKTITYQAVPSEVP NEPFKVATAYRPPALRNKPITNSKLHEEPPQ NMKPQSGNDKPLSKTALKNQRKHEAKKAAK QEARSDKSPDLAPITAPQSTPRNTVSQSISGDP EIDKKIKNLKKKLKAIEQLKEQAATGKQLEK NQLEKIQKETALLQELEDLELGI 384 1734 A 3242 3 678 IRSPAARSPGLETITCLIFVIAALAVFVDSAIP RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS ASLVRATVRAVSKRKLQFTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVFTRGCNGESYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEELLPEPGSETPTVASEALAELLHGALLRR GPEMGYLPOPPLGPEGGEETTTTITTTTT TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRFG ALATTSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGFQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVVLTYQ	1			1			PRNAAYYSPHGHILVLAGFGNLILQI*AD/IMK
NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP NEEPKVATAYRPPALRIKFITINSKLHEEEPPQ NMKPQSGNDKPLSKTALKNQRKHEAKKAAK QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP EIDKKIKNLKKKLKAEQLKEQAATGKQLEK NQLEKUQKETALLQELELGI 384 1734 A 3242 3 678 IRSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSFPYTILDPFLYLPGQCAPPQP LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS ASLVRATVAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEELLPPGSETFTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTT TVTSFVLCNNNISEGGGYVESPDLGSFVSRTL GLLDCTYSHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSFT NRLLHIFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLINGTRFSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDDNDRLMVRSGGSPLSPVIYDS DMDDVPBEGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATTSCLPGYALGLEPPGPPNAGCVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLPEDGGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPSGPNPGLGQGPVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ							VWNVKNYKLISKPVASDSTYFAWCPDGEHIL
NEPRVATAYRPPALRNKPITNSKLHEEPPPQ NMKPQSGNDKPLSKTALKNQRKHEAKKAK QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP EIDKKIKNLKKLKAIEQLKEQAATGKQLEK NQLEKIQKETALLQELEDLELGI 384 1734 A 3242 3 678 IRSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLRPTFGPFHRGAGTAKMS ASLVRATVRAVSKRILQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPQLLFLILLSCFWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPLGPEGGEEETTTTIITTTTTT TVTSPVLCNNNISEGEGVVESPDLGSPVSRTL GLLDCTYSIHVYGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSFT NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLINGTRPSWNGETPSCMASCGGTHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNIRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATTSCLPGYALEPPGPPNATECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLPGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPPPPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ	•		1	Į,	į		
NMKPQSGNDKPLSKTALKNQRKHEAKKAAR QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP EIDKKIKNLKKKLKAEQLKEQAATGKQLEK NQLEKIQKETALLQELEDLELGI 384 1734 A 3242 3 678 IRSPARSPGLETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS ASLVRATVRAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTVTT TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSHVYPGYGGIEQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLLHFQSRPVPGGGFRHYQAYLLSCFFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPRGLISDAQSLYVELLSETFANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRFG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLOFQAPPGPPRPGLGGGFVLHFKEVPR NDTCPFLPPPEWGWRTASHGDLIRGTVLTYQ							
QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP EIDKKIKNLKKKLKAIEQLKEQAATGKQLEK NQLEKIQKETALLQELEDLELGI 384 1734 A 3242 3 678 IRSPARSFGLETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSPPYTLLDPPLYLPGQCAPPQP LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS ASLVRATVRAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEETTTTIITTTTVTT TVTSPVLCNNNISEGGGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLANSSMLGEGQVLRSPT NRILLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMWRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAFFLAHGNVTTTDPEYRFG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPPACKAMCGGELSERAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPPPGGGGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ			1	1			•
BİDKKIKNLKKILAGÜLEQAATGRQLEK NQLEKIQKETALLQELEDLELGI 384 1734 A 3242 3 G78 IRSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS ASLVRATVRAVSKRKLQPTRAALTLTPSAVN KIKQLIKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEBILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPLQBFGGEETTTTIITTTVTI TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRILLHFQSPRVPRGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTHNA TLGRIVSPEPGAVGPNLTCRWVIEAAEGRRI. HLHFERVSLDEDNDRLMVRSGGSPLSPVYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPPGPPNAIECVDPTEPHW NDTEPPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFPGDFSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPPPPGGQFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ							
NQLEKIQKETALLQELEDLELGI	1			1			
1734 A 3242 3 678 IRSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLRPTFGPRHRGAGTAKMS ASLVRATVRAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTILL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI							
RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS ASLVRATVRAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTTVT TVTSPYLCNNNISEGGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLTGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ	384	1734	A	3242	3	678	
LSQČARRVHGEKLRRPTFGPRHRGÅGTAKMS ASLVRATVRAVSKRKLQPTRAALTLTPSAVN KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTUITTTTVTT TVTSPVLCNNNISEGGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLHFGSPRVPRGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ	1 50.		''		ļ [~]	0.0	
KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL EYTKTKGDSDEEVIQDGVRVPIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEETTTTIITTTTVTT TVTSPVLCNNNISEGGGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRILLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLVVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ				Į.			
EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTI TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ]		ļ		ASLVRATVRAVSKRKLQPTRAALTLTPSAVN
GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTT TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNIDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPPPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ							KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL
FNI 385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTT TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ							
385 1735 A 3243 3190 664 VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTT TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ					1	•	
KEEEILPEPGSETPTVASEALAELLHGALLRR GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTT TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ	205	1725	 _	2242	3100	661	
GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTT TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ	202	1/33	A	3243	3190	004	
TVTSPVLCNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ				1			
GLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELL VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ	1]]				
VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ			ł			Į	
NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP PRPAHGDVSVTDLHPGGTATFHCDSGYQLQG EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ			1				
EETLICLNGTRPSWNGETPSCMASCGGTIHNA TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ	1			1	}		NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP
TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ	1			1			
HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ	1		1	[T ₁
DMDDVPERGLISDAQSLYVELLSETPANPLLL SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ	1	ł	ļ	} .	ļ]	
SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ			1				
ALATFSCLPGYALEPPGPPNAIECVDPTEPHW NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ							
NDTEPACKAMCGGELSEPAGVVLSPDWPQS YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ							
YSPGQDCVWGVHVQEEKRILLQVEILNVREG DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ	1		1	1	}	l	
DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ				1			
GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ	1	ļ	l	1	}	1	
NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ]			GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR
CEPGYELLGSDILTCQWDLSWSAAPPACQKI							NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ
			<u></u>	<u> </u>	L	<u> </u>	CEPGYELLGSDILTCQWDLSWSAAPPACQKI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion MTCADPGEIANGHRTASDAGFPVGSHVQYRC LPGYSLEGAAMLTCYSRDTGTPKWSDRVPKC
						ALKYEPCLNPGVPENGYQTLYKHHYQAGESL RFFCYEGFELIGEVTITCVPGHPSQWTSQPPLC KVTQTTDPSRQLEGGNLALAILLPLGLVIVLG SGVYIYYTKLQGKSLFGFSGSHSYSPITVESDF SNPLYEAGDTREYEVSI
386	1736	A	3250	5725	3984	GTSTVTMATKKHFSIILNLLGMLLKKDNQDT RKLLMTWALEVAVVMKKSETYAPLFCLPSF HKFCKGLLADTLVEDVNICLQACSSLHALSSS LPDDLLQRCVDVCRVQLVHRGTCIRQAFGKL LKSIPLGVFLSNNNHTEIQEISLALRSHMSKAP SNTFHPQDFSD/VISFILYGNSHRTGKDNWLE RLFYSCQRLDKRDQSTIPRNLLKTDAVLWQW AIWEAAQFTVLSKLRTPLGRAQDTFQTIEGIIR SLAGHTLNPDQDVSQWTTADNDEGHGNNQL RLVLLLQYLENLEKLMYNAYEGCANALTSPP KVIRTFLYTNRQTCQDWLTRIRLSIMRVGLLA GQPAVTVRHGFDLLTEMKTTSLSQGNELEVSI MMVVEALCELHCPEAIQGIAVWSSSIVGKHL LWINSVAQQAEGRFEKASVEYQEHLCAMTG VDCCISSFDKSVLTLASAGCKSASLKHCLNGE SRKSVLSKPTDSSPEVINYLGNKACECYISTA DWAAVQEWQNAIHDLKKSTSSTSLNLKADF NYIKSLSSFESGKFVECTEQLELLPGENINLLA GGSKEKIDMKKLLRNM
387	1737	A	3255	380	76	MDIFLYNCKYQVQTEI*NSIQHIMA\SKKLSRF LKYVHNL*AENYKTLMK*INEDLNKQRDVPY S*TARLNKMSIPTKTIFRFKAIYIKIPATYFIET NMQ
388	1738	A	3260	685	428	PQWLGLQVYALPPANFVFFVEMRSTILAQTG FELLDSSDLPASASKSAGITCMSHHARTLSLK *WPFCLSATQEKFC*PASEGVAW
389	1739	A	3269	1	332	LDGYHTPIYMLNRIIRLPAAL*IISDQTGHALTI LTRLETQMINADYQNKLTLDYLLTTDREVYE PFNLTNYCLHIHNQRLGAYDLG*V*Q/KLAHV PVOV*HGFDPEAMFR
390	1740	A	3270	2	372	GRCHDONKGKS\DGPDAQAEACGGESTYQEL LVNQNPIGQPLACRRLTRKIYEGIKKAVKPNH SPRGVKKVHKFVNKGEKGIMVLAGDTLGIGV YCLLPCMC*DRKLTYAHIPSTTDLGAGAGY
391	1741	A	3273	1	187	FFQEMLDIMKAISDMMGKCTYPVLKEDAPRQ HVETFFQ\EELTRSQEGMKLGENFLMFAMPP DDSKESKGK*FFQEMLDIMKAISDMMGKCTY PVLKEDAPRQHVETFFQVGINQKSRGHEVRR KFPDVCHAPR
392	1742	A	3281	901	521	FFFGDGVSPCRQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVLPRQANFCIF/M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFALKGCLPRQKEGGTLNLI
393	1743	A	3283	385	3	RNRSVVPEFVLLGLSAGPQTQTLLFVLFVVIC LLTVMGNLLLLVVINADSCLHTPMYFFLGQL SFLDLCHSSVTAPKLLENLLSEKKTISVEGCM A*VFFVFATGGTESSLLAVMAYDRYVAIRTR G
394	1744	Ā	3284	575	1054	CTKCKADCDTCFNKNFCTKCKSGFYLHLGKC LDNCPEGLEANNHTMECVSIVHCEVSEWNP WSPCTKKGKTCGFKRGTETRVREIIQHPSAKG NLCPPTNETRKCTVQRKKCQKGERGKKGRE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RKRKKPNKGESKEAIPDSKSLESSKEIPEQREN KQQQ RVLYVPSMGFCILVAHGWQKISTKSVFKKLS
206	1016		2002			WICLSMVILTHSLKTFHRNWDWESEYTLFMS ALKVNKNNAKLWNNVGHALENEKNFERAL KYFLQATHVQPDDIGAHMNVGR
396	1746	A	3293	1	172	GFRAVVMTVKTEAAKGTLTYSRMRGMVAIL IAFMKQRRMGLNDFIQKIANNSYACKQ
397	1747	A	3295	12	401	AEPACGASSCTPPSLRSSSSQSVGPLRPGRPL WSEACAFL*AAAPQGPASPCCGLPSGFPRVW AQCCPPGGALRFPEGLGSVLSPRRCPQVSRGS GLSAVPQEVPSGFLGPGLRACPQEAPSRFLRA GLT
398	1748	A	3300	1912	2768	KQRRWQNIQRKGPKRYIVIAGNSQSHQPMIFS MLRKLPKVTCRDVLPEIRAICIEEIGCWMQSY STSFLTDSYLKYIGWTLHDKHREVRVKCVKA LKGLYGNRDLTARLELFTGRFKDWMVSMIV DREYSVAVEAVRLLILILKNMEGVLMDVDCE SVYPIV*ASN*GLASAVGEFLYWKLFYPECEI RTMGGREQRQSPGAQRTFFQLLLSFFVESKSH SVTQAGVQWQFSAHRDLCLPGSSNSHVSASR VAGIAGAHRHTWLIYVFFSWRQGFAVLAGL VSNS
399	1749	A	3301	536	2391	LRSYGCKAPSRISHLHK\FLELLPSLLMGYSE SPPPITDSWAPFISLTHHVLSQSQSPLSSNCWI CLSTHTQ*FTALPADLLTWTQSNVSLHISYLAI PFLADSFLKPV/L*PGNSAKHLSFKLSSLSMVS GRAVALLHLIASGLTSIQTNTASSKPPIWGY\L STQTSFISPPPLCLSRTYPNPAHATMVGQVPQ SLCGLIFTL/RTPCRPSILHPNYKIISTSAWQKV LCFSGSPTIHTSLHLTTGSSFLSFHPIPGFPAAN SALYVSSLKGPPGKNVTIPSPVTGT*QPPHRGS N/RLTVDKDNFFLSPKPNSLHQLPSQ\TPYQAL TGAALAGSYPIWENENTLSWLPTFTYNFCLST PSLFFLCDTN*YLCLPANWSGTCTLVFQAPTI NILPPNQTILISVEASISSSPIRNKWALHLITLLT GLGITAALGTGIAGITTSITSYQTLFTTLSNTVE DMHTSITSLQRQLDFLVGVILQNWRVLDLLT TEKGGTCIYLQEECCFCVNESGIVHIAVRRLH DRAAEL*HQVADSWWQGSSLLRWIPWVAPF LGPLIFLFLLLMIGPCIFNLVSRFISQRLNCFIQ ASMQKHIDNIFHLCHV*YQSLRGNHSEAPEPR P
400	1750	A	3303	2	453	THWRHSSGVPGSTTARRRRRELEIATSDNQE YYNRLCQEVTNRERNDQKMLADLDDLNRTK KYLEERLIELLRDKDALWQKSDALEFQQKLS AEERWLGDTEANHCLDCKREFSWMVRRHHC RICGRIFCYYCCNNYVLSKHGGKKERCC
401	1751	A	3304	1	626	MAPQHSSLDDKVPQQASTVCFEFQDILQHSQ CTEHKDSLWGPGARSQPFGAHNTRLSPDSCP EKIVLRALKDSRAGMPEQDKDPGVQENPDD QRRVPQGTGDAPSAFRPLWDNGGLSPFVSRP GPLERDLHAQRSEVTYNQRSQSSWMSSFPKR NAFVSPYSSMGQAQP/GLPKTNPIGESCCWEG LSLSTQILG*QKPSKYIPSLCKR
402	1752	A	3305	1678	172	MELPSGPGPERLFDSHRLPGDCFLLLVLLLYA PVGFCLLVLRLFLGIHVFLVSCALPDSVLRRF VVRTMCAVLGLVARQEDSGLRDHSVRVLISN HVTPFDHNIVNLLTTCSTVSESEAESATGRFP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GAQLKAPLSPLAFRMEDTEALPLTPILYPTCQ FFFF\IFLNIFLLAFSSPGSQPLLNSPPSFVCWSR GFMEMNGRGELVESLKRFCASTRLPPTPLLLF PEEEATNGREGLLRFSSWPFSIQDVVQPLTLQ VQRTLVSVTVSDASWVSELL\WSLFVPFTVY QVRWLRPVHRQLGEANEEFALRVQQ\LVAKE LG\QTGTRLTPA\DKAEHMKRQRHPR\LRPQS AQSSFPSPWVLSS/SDVQTGQTLGFREFKESF CPHVAIGVFIPERPWPKTGCCKTLTHLILLL*G GPVSFSCPE\DIHPRGT*VPTQQASGLPSFFSYG RABGGVI **UPS AQDATE TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACDATE ACCOUNT TEAUSCS WAR ACCOUNT TEAUSC WAR ACCOUNT TEAUSCS WAR ACCOUNT TEAUSCS WAR ACCOUNT TEAUSCS WAR ACCOUNT TEAUSCS WAR ACCOUNT TEAUSCS WAR ACCOUNT TEAUSCS WAR ACCOUNT TEAUSCS WAR ACCOUNT TEAUSCS WAR ACCOUNT TEAUSCS WAR ACCOUNT TEAUSCS WAR ACCOUNT TEAUSC WAR ACCOUNT TEAUSC WAR ACCOUNT TEAUSC WAR ACCOUNT TEAUSC WAR ACCOUNT TEAUSC WAR ACCOUNT TEAUSC WAR A
	1			!		PARGGVL*HPSAQQPLTFA\KSS\WARAGRAL QERKQ\ALYEYARRFTERRAPGGLD
403	1753	A	3307	44	447	DPSPSLLAVALGLRAGERTRSGPGSSSPSGGIS GGASAGLASSPECACGRSHFTCAVSALGECT CIPAQWQCDGDNDCGDHSDEDGCILPTCSPL DFHCDNGKCIRRSWVCDSDNDCEDDSDEQD CPPRECEED
404	1754	A	3311	409	1	PRHGWGRRVLGRDRPRLQKVKKSVKAIYIPG QDHVQNEEIYARVLDKFGSNFLSRDNADLGT AFVKFSTLTK*LSALLKNLLQGLSRNVIFTLDS LLKGDLKGVKGDLKKPFDKAWKDYETKFAK IEKEKREREWR
405	1755	A	3322	12	458	AAVPVENPWDDPRVRPRVRIFTWEDCIAGQA KVLCNDSYGVTIDWSPKGAFIRLTSQSVGNG HPASKENDQMVDTIKNTTKVPIIWTYGDMVE PRPQMIRPAVGAKHKELWKILMALKKIKVIWE GKYTKPSQYNPNYMLELAHNDSVW
406	1756	A	3324	1	426	LSMLSTISTEHRLSVLWPIWYCCHCPTHLSAV MCVLLWALSLLQSILEWMFCSFLFSDVDSDN WCQILDFLTAVWLIFLILVLCGFTLVLLVRIIC GSQKMPLTRLYVTILLTGLVFLFCSLPLSIQ*F LLYWIEKDLDDL
407	1757	A	3328	213	1841	SGDLSPAELMMLTIGDVIKQLIEAHEQGKDID LNKVKTKTAAKYGLSAQPRLVDIIAAVPPQY RKVLMPKLKAKPIRTASGIAVVAVMCKPHRC PHISFTGNICVYCPGGPDSDFEYSTQSYTGYEP TSMRAIRARYDPFLQTRHRIEQLKQLGHSVD KVEFIVMGGTFMALPEEYRDYFIRNLHDALS GHTSNNIYEAVKYSERSLTKCIGITIETRPDYC MKRHLSDMLTYGCTRLEIGVQSVYEDVARD TNRGHTVKAVCESFHLAKDSGFKVVAHMMP DLPNVGLERDIEQFTEFFENPAFRPDGLKLYP TLVIRGTGLYELWKSGRYKSYSPSDLVELVA RILALVPPWTRVYRVQRDIPMPLVSSGVEHG NLRELALARMKDLGIQCRDVRTREVGIQEIH HKVRPYQVELVRRDYVANGGWETFLSYEDP DQDILIGLLRLRKCSEETFRFELGGGVSIVREL HVYGSVVPVSSRDPTKFQHQGFGMLLMEEA ERIAREEHGSGKIAVISGVGTRNYYRKIGYRL QGPYMVKMLK
408	1758	A	3335	3	467	AIASPRAAGIRHELTSTMAAGKNKRLTKGGK KGAKKKAV/DNIINIGKTLVTRTQRTKIASDG LKGRVFEESLADLQND\TDGYLLRVI*VAFTT ERTNQI/REVFNKLIPDSIGKDIEKACQSIYPLH DDFARKVKMLKKPKFELRKLMELHGEGSS
409	1759	A	3338	7	1252	PRWRNSARDEILLSFPQNYYIQWLNGSLIHGL WNLASLFSNLCLFVLMPFAFFFLESEGFAGLK KGIRARILETLGMLLLLALLILGIVWVASALID NDAASMESLYDLWEFYLPYLYSCISLMGCLL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLLCTPVGL\SRMFTVMGQLLVKPTILEDLDE QIYIITLEEEALQRPTKWAVFIRW/KYNIMELE QELENVKTLKTKLERRKKASAWERNLVYPA
						VMVLLLIETSISVLLVACNILCLLVDETAMPK GTRGPGIGNASLSTFGFVGAALEIILIFYLMVS SVVGFYSLRFFGNFTPKKDDTTMTKIIGNCVS ILVLSSALPVMSRTLGITRFDLLGDFGRFNWL GNFYIVLSYNLLFAIVTTLCLVRKFTSAVREE LFKALGLHKLHLPNTSRDSETAKPSVNGHQK AL
410	1760	A	3339	127	1433	GSHRFSLASPLDPEVGPYCDTPTMRTLFNLL WLALACSPVHTTLSKSDAKKAASKTLLEKSQ FSDKPVQDRGLVVTDLKAESVVLEHRSYCSA KARDRHFAGDVLGYVTPWNSHGYDVTKVFG SKFTQISPVWLQLKRRGREMFEVTGLHDVDQ GWMRAVRKHAKGLP*CLGSCLRTGLTMISG/ YVLDSEDEIEELSKTVVQVAKNQHFDGFVVE VWNQLLSQKRVGLIHMLTHLAEALHQARLL ALLVIPPAITPGTDQLGMFTHKEFEQLAPVLD GFSLMTYDYSTAHQPGPNAPLSWVRACVQV LDPKSKWRSKILLGLNFYGMDYATSKDAREP VVGARYIQTLKDHRPRMVWDSQVSEHFFEY KKSRSGRHVVFYPTLKSLQVRLELARELGVG VSIWELGQGLDYFYDLL*VGIAASAVDVFFSK PWSE
411	1761	A	3342		2701	VATRKLAKGFTQFAKMTEGTKKTSKKFKFK FKGFGSFSNLPRSFTLRRSSASISRQSHLEPDTF EATQDDMYTVPKSPPAYARSSDMYSHMGTM PRPSIKKAQNSQAARQAQEAGPKPNLVPGGV PDPPGLEAAKEVMVKATGPLEDTPAMEPNPS AVEVDPIRKPEVPTGDVEEERPPRDVHSERAA GEPEAGSDYVKFSKEKYILDSSPEKLHKELEE ELKLSSTDLRSHAWYHGRIPREVSETLVQRN GDFLIRDSLTSLGDYVLTCRWRNQALHFKIN KVVVKAGESYTHIQYLFEQESFDHVPALVRY HVGSRKAVSEQSGAIIYCPVNRTFPLRYLEAS YGLGQGSSKPASPVSPSGPKGSHMKRRSVTM TDGLTADKVTRSDGCPTSTSLPRPDSIRSCA LSMDQIPDLHSPMSPISESPSSPAYSTVTRVHA APAAPSATALPASPVARRSSEPQLCPGSAPKT HGESDKGPHTSPSHTLGKASPSPSLSSYSDPDS GHYCQLQPVRGSREWAATETSSQQARSYGE RLKELSENGAPEGDWGKTFTVPIVEVTSSFNP ATFQSLLIPRDNRPLEVGLLRKVKELLAEVDA RTLARHVTKVDCLVARILGVTKEMQTLMGV RWGMELLTLPHGRKLRLDLLERFHTMSIML AVDILGCTGSAEERAALLHKTIQLAAELRGT MGNMFSFAAVMGALDMAQISRLEQTWVTLR QRHTEGAILYEKKLKPFLKSLNEGKEGPPLSN TTFPHVLPLITLLECDSAPPEGPEPWGSTEHGV EVVLAHLEAARTVAHHGGLYHTNAEVKLQG FQARPELLEVFSTEFQMRLLWGSQGASSSQA RRYEKFDKVLTALSHKLEPAVRSSEL
412	1762	A	3347	1	898	IDRAAECRTKPLPMAVSIRGNADSIVACLVLM VLYLIKKRLVACAAVFYGFAVHMKIYPETYI LPITLHLLPDRDNDKSLRQFRYTFQACL*ELL KRLCNRTALMFVAVAGLTFFALSFGFYYEYG WEFLEHTYFYHLTRRDIRHNFSPYFYMLYLT AESKWSFSLGIAAFLPQLILLSAVSFAYYRDL VFCWFLHTSIFVTFNKVCTSQYFLWYLCLLPL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VMPLVRMPWKRAVVLLMLWFIGQAMWLAP AYVLEFQGKNTFLFIWLAGLFFLLINCSILIQII SHYKEEPLTERIKYD PIPVRWNSLEGRLLRGYEQHANDGKDYISRN
413	1,03	A	3301	3	4/4	*DLRSWTAADMAAQITKRKWEAEEFAEQIKA YLEGTCVER/LRTHLENGKETLQLTEQSSQPTI PIVGIVAGLVLLGAVVTGAVVSAVMCRKKNS GHFLPTDRVSYSEAASSDHAQGSDVSLTACK V
414	1764	A	3363	1488	453	HQILELKKKILKTYNPDYDEDLVQEASSEDVL GVHMVDKDTERDIEMKRQLRRLRELHLYST WKKYQEAMKTSLGVPQRERDEGSLGKPLCP PEILSETLPGSVKKRVCFPSEDHLEEFIAEHLP EASNQSLLTVAHADAGTQTNGDLEDLEEHGP GQTVSEEATEVHMMEGDPDTLAELLIRDVLQ ELSSYNGEEE\DPEEVKTSLGVPQRGDLEDLE EHVPGQTVSEEATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEEATGVHMMQVDPATLA KQLEDSTITGSHQQMSASPSSAPAEEATEKTK VEEEVKTRKPKKKTRKPSKKSRWNVLKCWD IFNIF
415	1765	A	3369	431	315	IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP PSFSPSL
416	1766	A	3373	42	651	RQEKMGLGEIGASGVLRSMLKERKKONMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSPLHPLQPQQEASTCPSGTLQGREKAA PGQGRPLCSLWAGGAGA\PGERGAEGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH
417	1767	A	3382	2	2061	EAQDPRACGPDAGGRFAARDAPGNSLRPPPS SPP/GWPGQLRLLPRVPGSELRCGKPERGRLP ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGA\RGPRLSLGPPPGPPPAPALPAR ASAGAGAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPGPQPGPPPMPARPR*AS\S TRGSRRGPGSRPARAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPQQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPITSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK
418	1768	A	3398	304	2121	EEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN NGKEESKIDHLTNNRNDLISKEEQNSSSLLEE

	000 10			-		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ĺ		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	l		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
		<u> </u>		sequence		nucleotide insertion
	1	1				NKVHADLVISKPVSKSPERLRKDIEVLSEDTD
		1		l		YEEDEVTKKRKDVKKDTTDKSSKPQIKRGKR
l .]	J	RYCNTEECLKTGSPGKKEEKAKNKESLCMEN
[[ĺ				SSNSSSDEDEEETKAKMTPTKKYNGLEEKRK
	ļ		İ	İ		SLRTTGFYSGFSEVAEKRIKLLNNSDERLQNS
	į					RAKDRKDVWSSIQGQWPKKTLKELFSDSDTE
]	AAASPPHPAPEEGVAEESLQTVAEEESCSPSV
	ŀ)		<u> </u>]	ELEKPPPVNVDSKPIEEKTVEVNDRKAEFPSS
		ł	1	1	1	GSNFSA*IPLPYLHLNRLHQSL*QKGSRQQSS
	ļ		į	1		VTVSEPLAPNQEEVRSIKSETDSTIEVDSVAGE
						LQDLQSERE*LASRF*CQCELKQ**SARTRTS*
		1			ļ	KSLYRSEKSERCSGRRKFIKKAEKKP*SNSGK
	L	Ì				QQKEGK
419	1769	Α	3399	206	463	QRECLSIHIGQAGIQIGDACWELYCLEHGIQP
	1			į		NGVVLDTQQDQLENAKMEHTNASFDTFFCE
	}	ļ	ļ			TRAGKHVPRALFVDLEPTVIDGIR
420	1770	A	3408	1010	685	RRLSFFF*IWSSVLVTQARVQWRDLGSPQPLP
				ł		PGFKRFSCLSLPSSWDYRHPSPRPVNF/HVFLV
		l				VMGFHHVGQAGLELLTSGDLPALASQSARIT
				i	}	GVNHCAQPRGHFH
421	1771	A	3409	355	1326	ADSNLIESCWQELGLGPWGGDWRVEQVGAS
	ļ		1			ASLRFPREVCSIRFLFTAVSLLSLFLSAFWLGL
		ł	1			LYLVSPLENEPKEMLTLSEYHERVRSQGQQL
ļ						QQLQAELDKLHKEVSTVRAANSERVAKLVF
		Į	i	1		QRLNEDFVRKPDYALSSVGASIDLQKTSHDY
	1	1	ŀ	1		ADRNTAYFWNRFSFWNYARPPTVILEPHVFP
			1	Į.		GNCWAFEGDQGQVVIQLPGRVQLSDITLQHP
	1	İ	1	1		PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ
	ļ.	ł			ļ	VYDETEVSLGKFTFDVEKSEIQTFHLQNDPPA
	j		Į]		AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT
		1				SEGAEGSÄQGPH
422	1772	A	3412	2	421	EFDAOPSIGALVVFKRP*ATTGSDPGPKRGMN
]		1] _		YLVSCSMRSPESGKGEPGTARDYTPMGRPPP
	1					PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG
		ļ		ļ		QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP
		1				VDTAGAPASPGPDVCE
423	1773	A	3420	91	706	DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/
1.2	****	1 **	1	1	1	RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG
			1	1	1	EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS
		1	1	1		KTHLPGFVEQAEALKAKGVQVVACLSVNDA
		1			1	FVTGEWGRAHKAEGKVRLLADPTGAFGKET
	1	1				DLLLDDSLVSIFGNRRLKRFSMVVODGIVKA
		1	İ	ĺ	1	LNVEPDGTGLTCSLAPNIISQL
424	1774	A	3421	4	7688	ROVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ
724	1//4	^	3421	1	'000	FSEKRYVVQVREDVTPGAPVLRVTASDRDKG
	}	1	l	.}	1	SNAVVHYSIMSGNARGQFYLDAQTGALDVV
		ŀ		ŀ		SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL
			1			VTVQVLDINDNAPIFVSTPFQATVLESVPLGY
	1		1			LVLHVQAIDADAGDNARLEYRLAGVGHDFP
	1	1.	1	1		FTINNGTGWISVAAELDREEVDFYSFGVEAR
1	1	1	1	1	Ì	
1		1				DHGTPALTASASVSVTALDVNDNNPTFTQPE YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY
-		1	1	}		
		1	1	1		QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE
		l	1	ļ		RQYVLAVTASDGTRQDTAQIVVNVTDANTH
1	1					RPVFQSSHYTVNVNEDRPAGTTVVLISATDE
				1	i .	DTGENARITYFMEDSIPQFRIDADTGAVTTQA
1						ELDYEDQVSYTLAITARDNGIPQKSDTTYLEI
1	1	1	1			LVNDVNDNAPQFLRDSYQGSVYEDVPPFTSV
1		1	I	1	1	LQISATDRDSGLNGRVFYTFQGGDDGDGDFI

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	"00	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	donoc		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Lacinos	!	1	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	l	ŀ	İ	residue of	Sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ŀ		peptide	Soquence	/=possible nucleotide deletion, \=possible
	1	ſ		sequence		nucleotide insertion
	 			sequence		1
		İ				VESTSGIVRTLRRLDRENVAQYVLRAYAVDK
1	·	}				GMPPARTPMEVTVTVLDVNDNPPVFEQDEFD
						VFVEENSPIGLAVARVTATDPDEGTNAQIMY
						QIVEGNIPEVFQLDIFSGELTALVDLDYEDRPE
	1					YVLVIQATSAPLVSRATVHVRLLDRNDNPPV
1	}	ľ	1			LGNFEILFNNYVTNRSSSFPGGAIGRVPAHDP
						DISDSLTYSFERGNELSLVLLNASTGELKLSR
						ALDNNRPLEAIMSVLVSDGVHSVTAQCALRV
1	}]				TIITDEMLTHSITLRLEDMSPERFLSPLLGLFIQ
						AVAATLATPPDHVVVFNVQRDTDAPGGHILN
						VSLSVGQPPGPGGGPPFLPSEDLQERLYLNRS
[,	LLTAISAQRVLPFDDNICLREPCENYMRCVSV
						LRFDSSAPFIASSSVLFRPIHPVGGLRCRCPPGF
						TGDYCETEVDLCYSRPCGPHGRCRSREGGYT
						CLCRDGYTGEHCEVSARSGRCTPGVCKNGGT
						CVNLLVGGFKCDCPSGDFEKPYCQVTTRSFP
						AHSFITFRGLRQRFHFTLALSFATKERDGLLL
1	·			_		YNGRFNEKHDFVALEVIQEQVQLTFSAGEST
1						TTVSPFVPGGVSDGQWHTVQLKYYNKPLLG
ļ						QTGLPQGPSEQKVAVVTVDGCDTGVALRFGS
		İ				VLGNYSCAA\QGTQGGSKKSLDLTGPLLLGG
1						VPDLPESFPVRMRQFVGCMRNLQVDSRHIDM
						ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC
<u> </u>						VNQWDAFSCECPLGFGGKSCAQEMANPQHF
1						LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD
	ľ					GVLLQAITRGRSTITLQLREGHVMLSVEGTGL
						QASSLRLEPGRANDGDWHHAQLALGAIGGP
						GHAILSFDYGQQRAEGNLGPRLHGLHLSNITV
i i						GGIPGPAGGVARGFRGCLQGVRVSDTPEGVN
						SLDPSHGESINVEQGCSLPDPCDSNPCPANSY
						CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC
						EHQSVCTRKPSAPHGYTCECPPNYLGPYCET
						RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC
						NKTSGECHCKENHYRPPGSPTCLLCDCYPTG
						SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF
						AEVTTNGCEVNYDSCPRAIEAGIWWPRTRFG
1						LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF
						NCTSITFSELKGFAERLQRNESGLDSGRSQQL
						ALLLRNATQHTAGYFGSDVKVAYQLATRLL
						AHESTQRGFGLSATQDVHFTENLLRVGSALL
						DTANKRHWELIQQTEGGTAWLLQHYEAYAS
						ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN
]						FAGAKLPRYEALRGEQPPDLETTVILPESVFR
						ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ
1						GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK
						RPIINTPVVSISVHDDEELLPRALDKPVTVOFR
						LLETEERTKPICVFWNHSILVSGTGGWSARGC
						EVVFRNESHVSCQCNHMTSFAVLMDVSRRE
[l	NGEILPLKTLTYVALGVTLAALLLTFFFLTLL
						RILRSNQHGIRRNLTAALGLAQLVFLLGINQA
						DLPFACTVIAILLHFLYLCTFSWALLEALHLY
					· ,	RALTEVRDVNTGPMRFYYMLGWGVPAFITG
				·		LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP
					ľ	VAFAVSMSVFLYILAARASCAAQRQGFEKKG
]						PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL
				i		FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK
					1	LACSRKPSPDPALTTKSTLTSSYNCPSPYADG
[' '						RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR
1						
					·	EESALNPG\OGPPGLGGIDGDA CET CDEKTOO
1						EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ H\DS*TRDFDSDLSLEDDQSGSYASTHSSDSEE

CEC III	I CEO ID	Met	SEQ	Desdisted	Deadisted and	L Amino coid accusage (A—Alenino C—Custaine
SEQ ID NO: of	SEQ ID NO: of	hod	ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid. E=Glutamic Acid.
-	1	400		beginning nucleotide	nucleotide	
nucl-	peptide	ļ	in		location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	i	Į	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ľ	1	Í		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
['		i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
			1	sequence		nucleotide insertion
		 			 	EEEEEEEAAFPGEQGWDSLLGPGAERLPLHS
ļ	1	į	l	ļ		TPKDGGPGPGKAPWPGDFGTTAKESSGNGAP
		l	ļ			EERLRENGDALSREGSLGPLPGSSAQPHKGIL
		1	l			
1			ł			KKKCLPTISEKSSLLRLPLEQCTGSSRGSSASE
1			İ		۱.	GSRGGPPSRPPPRQSLQEQLNGVMPIAMSIKA
Ĺ	<u> </u>	L				GTVDEDSSGSEFLFFNFLH
425	1775	A	3429	155	1417	GEPAVQSCDCGCTQRSCPWLLVAPGLLSSSSS
			1			RAASVREAEDAPLQPASIHPVSQGSRGPEGSL
			1			GSAECLPGDPLGARRATRAHSPVPGPPPSLPA
			Į.			AGTAVKRGLQPG*GA/GATSTPGTGAATGGL
			ļ			CGPAWAAPSAVGPCCCCPSISTTPSQMRSARP
}	ł	l	ł			
1	1	l	l .			SLGCLPSWAS\PGTEHPPGPQGPGPS*DLCSV*
1	1	l				KREFQRGPWAGMVILHRISAADPARAPGPDS
			1			NLQSALQQPATGCSEPAAVYSPPIGLWGA**P
		1]			EYG*PQHSLPG*TAPADR*P\AGIKDRVYSNSI
1	1		1			YELLENGQRAGTCVLEYATPLQTLFAMSQYS
1			l			QAGFSREDRLEQAKLFCRTLEDILADAPESQN
						NCRLIAYQEPADDSSFSLSQEVLRHLRQEEKE
	}				·	EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP
ļ	1	l	ļ	ļ		LPLRTDFS
426	1776	A	3431	1662	369	AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR
420	1770	^	3731	1002	303	SSQSLHCLSPRQPNHYLQALRAVGGICQDYD/
İ						
						SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP
			Į.			AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE
	1]	j			KAGPHCSRLALTG\SHDFAINFDPENPECEGK
	1					RGDFHLPRLPADTLHTGAQTPLPRAQLPVPST
						HPRPVFI\EISGVIASYRRCLPQIQLYGPTNVAP
	ļ					INRVAEPAQREQSTGQATKYSVLLVLTDGV
	l		1			VSDMAETRTAIVRASRLPMSIIIVGVGNADFS
İ	1	ł	ì			DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR
						DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD
	1					VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI
		Ì				TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ
100	1333	4:	2446		0540	GISPGAPRPCTLATTPSPSP
427	1777	A.	3446 .	79	9748	GCQSCWPAWPRLRRRGPASAGARLGRKAPW
		i				GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA
1		1		ļ		ASRPEASGDCRAGRETAMATLEKLMKAFESL
1		1		'		KSFQQQQQQQQQQQQQQQQQQQPPPP
		J				PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPPPPPPPP
1	ŀ					GPAVAEEPLHRPKKELSATKKDRVNHCLTIC
İ		1				ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA
						ESDVRMVADECLNKVIKALMDSNLPRLQLEL
	1	ļ				YKEIKKNGAPRSLRAALWRFAELAHLVRPQK
1	i	l		!		CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP
1						KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI
1	ļ					RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG
	l					LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV
	1	l	'			KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL
1	1	1		1		TLHHTQHQDHNVVTGALELLQQLFRTPPPEL
1	1]				LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI
1		1				AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS
1		1				ESRSDVSSSALTASVKDEISGELAASSGVSTPG
	<u> </u>					SAGHDITTEOPRSONTLOADSVDLASCDLTSS
	1	[(ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG
						TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD
			,			CTDMOVE OF OTODODODODO A TOT DOD A TO
1						GTDNQYLGLQIGQPQDEDEEATGILPDEASEA
1		1				FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF
1	1	l				VLRDEATEPGDQENKPCRIKGDIGQSTDDDS
İ		1	1			APLVHCVRLLSASFLLTGGKNVLVPDRDVRV
L	1	1				SVKALALSCVGAAVALHPESFFSKLYKVPLD
		·	•			

uence u	SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
seq- uence 914 914 916 1914 916 1914 916 1914 916 916	_	1					
uence 9149 got orrespond gl ist amino gl if ist amino gl if ist gl ist	eotide	seq-		USSN		corresponding	
mino acid reduce of peptide requence peptide sequence Talinconine, Va-Vaine, Wa-Tyuptophan, Ya-Tyusine, X-Unknow, Ya-Sup codon, Peposible muclocide deletion, ∀posible muclocide deletione	seq-	uence		09/496	correspondi		M=Methionine, N=Asparagine, P=Proline,
residue of peptide sequence Poptide mucleotide insertion Report of the mucleotide insertion investable nucleotide insertion Report of the mucleotide insertion investable nucleotide insertion Report of the mucleotide insertion in the mucleotide insert	uence			914			
peptide sequence Possible mueleotide deletion, =possible	1						
Invelocité insertion	1					sequence	
TIEVPEEQYNDILMYDHODPQURGATAILC GTLICELSRSRPHVOROMOTIFILINGNITS. ADCIPLLRKITLKDESSVTCKLACTAVRNCYM SLCSSYSPELGIQLIDVILTIRNSSYWJ.NTFEL LETLAEIDFELVSFLEAKAENLIRGAHHYTGI. LKLOEVLNNVVHILLGDEPPRRIFYAAAS. RIVPKLEYKCDGQADPVAVARDQSSVYL KLAHEIDFOPSHSYSTITTRYGAYHLIPSTD VTMENNLSRVIAAVSHELITSTTRALTFOCCE ALGLISTAFPVCWSLGWHGOVPPLSASDESR KSCTVGMATMILTLSSAWFIELSASDESR KSCTVGMATMILTLSSAWFIELSASDESR KSCTVGMATMILTLSSAWFIELSAHDAAL LAGNILAASAPKLSSWASEEANPAATK QEEVWPALGDRALVBWIZGLFSHLLKVINIC AHVLDDVAPCPAIRAKALPSITNPYSLSPIRK GKEKEPGCASVPLSFKKGEASAASRQCDTS GPYTTIKSSASSTYHLFSTLKLIDVLKATHA WYKYTLDLQNSTERFGGRISSALDVLSQILEL ATIQDICKCVEELGYHKSCTSRPMMATVC VQLLKTLFGTNLASGFGRISSALDVLSQILEL ATIQDICKCVEELGYHKSCTSRPMMATVC VQLLKTLFGTNLASGFGRISSALDVLSQILEL ATIQDICKCVEELGYHKSCTSRPMMATVC VQLLKTLFGTNLASGFGRISSALDVLSQILEL ATIQDICKCVEELGYHKSCTSRPMMATVC VQLLKTLFGTNLASGFGRISSALDVLSQILEL ATIQDICKCVEELGYHKSCTSRPMMATVC VQLLKTLFGTNLASGFGRISSALDVLSQILEL ATIQDICKCVEELGYHKSCTSRPMMATVC VQLLKTLFGTNLASGFGRISSALDVLSQILEL ATIQDICKCVEELGYHKSCTSRPMMATVC VQLKTLFGTNLASGFFYRDMATVC VQLKTLFGTNLASGFFYRDMATVC VQLKTLFGTNLASGFFYRDMATVC VQLKTLFFTNLASGFFYRDMATVC VQLKTLFFTNLAGGRISSAGRA QRLGSSSYRGGLHYCTCHAAPVCLL BSDQVFIGFVLKQFFYEVGGFRSEAHDNIFF FRYSTSYRGGRISSAGRA QRLGSSYRGGLHYCTCHAAPVCLL BSDQVFIGFVLKQFFYEVGGFRSEAHDNIFF FRYSTSYRFSKNGMATPKILLQUQUKTCHAAPVCLL BSDQVFIGFVLKQFFYEVGGFRSEAHDNIFF FRYSTSYRTSKNGMASGRIFTNLAGGRISSAGRA KAVTHAIPALQPVINDLFVLKGTTNKARDAGKE LETYKEVVSNLLRILRFSSLRVFOMLISSAGR KAVTHAIPALQPVINDLFVLKGTTNKARDAGKE LETYKEVVSNLLRILRFSSLRVFOMLISSAGR VTYPHTMASVSTVQLWISGILAHRVLSQSTED JVLSRIGHESSRYFLLSCTVINRLRGDGSTSTILE BHSEGGRINDERSSRYFLISCTVINRLRGDGSTSTILE BHSEGGRINDERSSRYFLOCTVINRLRGDGSTSTILE BHSEGGRINDERSSRYFLOCTVINRLRGDGSTSTILE BHSEGGRINDERSSRYFLOCTVINRLRGDGSTSTILE BHSEGGRINDERSSRYFLOCTHARDSSRYFOMLISSAGE DIVLSRIGHESSRYFLLSCHARDSS QCHTRSDSALLEGGSTTLDSLNLM ARSMITTHPALVLLWCQLLLVNHTDYRWW ARVQQTFYKHALLSTLYVCHLIBEROFT JVLNSRIGHESSRYFLLSTYPHKWTLOCHTONKYT THAACEMVABMASGLAFFA AREVTLARVSGTVQLGCCLALQFOL WSVYSTEPVTHACSLITYCHPREX JVLWSLESHHHLHPERKEDINKFVVATLEAL SWHLIBEGGRISTGGNSALFFA AREVTLA							, · ·
GTLICSILSRSRPHVGDWMGTIRTLTGNTPSL ADCPILLEKTLADESSYTCKLACTAVENCOM SLCSSSYSELGLOLIDVILTLRNSSYWLVRTEL LETLALDIFELLSTLAGEDFELLSTLAGENTYAAASL IRLVPKLETYCDGGGADPVAVAARDGSSYVL KLLMEETGPPSHSVSTTIRTYRGYNLLPSITD VTMENNLSSVIAAVSBEITSTTRALTFGCCE ALGLISTAFPVCWSLGWHGGVPPLSASDESR KSCTVGMATMILTLISAWPFDLDSASDESR KSCTVGMATMILTLISAWPFDLDSASDESR KSCTVGMATMILTLISAWPFDLDSASDESR KSCTVGMATMILTLISAWPFDLDSASDESR KSCTVGMATMILTLISAWPFDLDSASDESR KSCTVGMATMILTLISAWPFDLDSASDESR KSCTVGMATMILTLISAWPFDLDSASDESR KSCTVGMATMILTLISAWPFDLDSASDESR KSCTVGMATMILTLISAWPFDLDSASDESR KSCTVGMATMILTLISAWFDLDSASDESR KSCTVGMATMILTLISAWFDLDSASDESR KSCTVGMATMILTSINGPSFIRK GKEREFGEGASVLJSFKKGSBASASRGSDTS GPVTTSKSSSLGSFYHFSYLKJEDVLKATHA NYKVTLDLQNSTERFGGFLRSALDVLSQILEL ATLODIGKCVEELGYKLSCSFREPMATVC VQULKTLIFGTINLASQFDGLSSNPSKSQGRA QRLGSSSVRGCHYVCVKSCTSREPMATVC VQULKTLIFGTINLASQFDGLSSNPSKSQGRA QRLGSSSVRGCHYVCVKSCTSREPMATVC VQULKTLIFGTINLASQFDGLSSNPSKSQGRA QRLGSSSVRGCHYVCVKSCTSREPMATVC VQULKTLIFGTINLASQFDGLSSNPSKSQGRA QRLGSSSVRGCHYVCVKSCTSREPMATVC VQULKTLIFGTINLASQFDGLSSNPSKSQGRA QRLGSSVRGCHYVCVKSCTSREPMATVC VQULKTLIFGTINLASQFDGLSSNPSKSQGRA QRLGSSVRGCHYVCVGNATVCLL DSQVVTSNRADKNAHNIELJEELVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLKYNYCLL DSQVVTSNRADKNAHNIELJEELVIKALKQ YTTTTTAASQFTQULKSCSTED FLULSSRADGLAGENTYFTGLAKARA KAYHHAPALQPYHUDJFVLRGTKKARAOGNE LETTQKEVVVSNLIRLIQYHQVLEMFLVQQMH DSHEALGVLNTLFELLAPSSLRPVDMLLRSMF VTPITMASVSTVQLWSGILLARVLYLGSTED FLULSSRADGLAGENTYCQGLGTLLMCLHIFKS GMFRITAAATRIFRSFRLQLVGGILLEDIVT KQLKVEMSEQQHTFVCQGLILLMCLHIFKS GMFRITAAATRIFRSFRLQLVGGILLEDIVT KQLKVEMSEQGHTFVCQGLILLMCLHIFKS GMFRITAAATRIFRSFRLQLVGGLEDIVA ARLGMCREEVRRGALLIFCDYVCQNLHDES HTSGSKOKLYLFTSFRLLQLVGGLEDIVA ARLGMCREEVRRGALLIFCDYVCQNLHDES HTSGSKOKLYLTLYVBRILLSPRVOMBCLSSP PVSSSPLDGGGHVSLLTVPSPKDVWYLLVX SQCVTRSDEALLIFCDYVCQNLHDES HTSGSKALTLYVHIQDISLSPPVQDFSLAVRNN AASGLFIOAUGRGCHALTLYVDRILLSFNATAGAVT LVVXSLISHLHLHPERSEDIVKRVATALLQV LVVVSLISHLHHLPFEKEDIVKRVATALLQV LVVVSLISHLHLHPERSEDIVKRVATALLQV UVSSLESPPCEDTERTQNNLAVQGLT WSCHEFFTYNNLAVAVQGE GESPPEEDTERTQNNLAVAVQRISLVLSAMT VPAGMPAVSCLQGGPRKIPLALLDGNS					sequence		
ADCIPLIRATILADESSYTCKLACTAYRNOWN SLCSSSYSELGLQIDVUTLRNSSYN_VETEL LETLAEIDFRLVSTLEAKARNLRRGAHHYTCL LETLAEIDFRLVSTLEAKARNLRRGAHHYTCL LETLAEIDFRLVSTLEAKARNLRRGAHHYTCL LETLAEIDFRLVSTLEAKARNLRRGAHHYTCL LETLAEIDFRLVSTLEAKARNLRRGAHHYTCL LETLAEIDFRLVSTLEAKARNLRRGAHHYTCL LETLAEIDFRLVSTLEAKARNLRGAHHYTCL LETLAEIDFRLVSTLEAKARNLRGAHYTCL LETLAEIDFRLVSTLEAKARNLRGAHYTCL LETLAEIDFRLVSTLEAKARNLRGAHYTCH LETLAEIDFRLATHATTALTGCCC LETLAEIDFRLATHATTALTGCCC LACLLSTAFPVCUNSIGWICGVPLASADESR KSCTVGMATMITLLSSAWPPLDLSAHDQAL LAGNILLASAPKALSELNSWASEEBANPAATAL QEEVWPALGDRALVPMYSOLFSHLLKVINIC APPLIDDVAPGFAKAALPSLTTPPLASAPTAL QEEVWPALGDRALVPMYSOLFSHLLKVINIC APPLIDDVAPGFAKAALPSLTTPPLASHTAL GEKEPGGCASVPLSPKKGGEASAASRQSDTS GPYTTSKSSSLGSTYHLEPYLKLAUDUTACHTA NYKYTLDLQNSTEKFGGERSALDVLSQILE ATLQDICKCVEERIGVYLSCYSTEMATVC VOQLLKTLFOTNLASQFDGLSNPSKSQGRA QELGSSSVRFGLYHLCYMACHTPTOQLADA SLRNMVQAGDNDTSGWFDVLQKVSTOLKT NLTSWTKORADNAGHNBHRILFEPTVKALKQ VOLLKTLFOTNLASQFDGLSNPSKSQGRA QELGSSSVRFGLYHCYMACHTPTQALADA SLRNMVQAGDNDTSGWFDVLQKVSTOLKT NLTSWTKORADNAGHNBHRILFEPTVKALKQ NTTTTCVQLKQVLDLAGVQFRESEAIIPNEF FLVLLSYERSYNGMICHTROQLOGGMASGR KAVTHAPALGPWFDLFVLRGTNKADAGGR KAVTHAPALGPWFDLFVLRGTNKADAGGR KAVTHAPALGPWFDLFVLRGTNKADAGGR KAVTHAPALGPWFDLFVLRGTNKADAGGR KAVTHAPALGPWFDLFVLRGTNKADAGGR KAVTHAPALGPWFDLFVLRGTNKADAGGR VTPHTMASSYSTVQLWISGILAIRVLISQSTSD VISIGUSLSSFFTLISTTSNINLADONSTSTLE EHSBGRQGKNLPEETSSFLLQVGULLEDIVT RQLKVEMSEQQHTPKVGLGGGSFYTLDSLINL ARSWTTHALVLILGCTLIVNTTDYRWW ARSVQTTKKRRSLSSTALLSCTVNINLADDONSTSTLE EHSBGRQGKNLPEETSSFLLQVGULLEDIVT RQLKVEMSEQGHTPVCQLGCTSTYDMLKTLTQCLCGG HLSQSGAVLTLYVORLCTPRVLARMVDUL ARRAWGRADALAGALTLYVORLCTPRVLARMVDUL ARRAWGRADALAGALTLYVORLCTPRVLARMVDUL ARRAWGRADALAGARTHAGAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA		[
SLCSSYSPELGARAENHIRGAHNTGL LETLAEDFRIVSTEARAENHIRGAHNTGL LIKLQERVINNVVIHLLGDEDPRVRIVAARSL RILVPKLFYKCDQGADPVVAVARDQSSVYL KLLMEETQPFSIFSVSTITRYTRGYNLLPSITD VTMENNLSRVIAAVSHEITSTTRTGCCE ALCLLSTAFPVCIWSLGWIFGVPPLSASDESR KSCTVGMATMILTILSSAWFPLIGTGCE ALCLLSTAFPVCIWSLGWIFGVPPLSASDESR KSCTVGMATMILTILSSAWFPLIGTLKVINIC QEVVPALGDRALVPAVGQLESHLLVINIC AHVLDDVAPGPAKAALPSILLSVINIC AHVLDDVAPGPAKAALPSILLSVINIC AHVLDDVAPGPAKAALPSILLSVINIC AHVLDDVAPGPAKAALPSILLSVINIC AHVLDDVAFGPAKAALPSILLSVINIC AHVLDDVAFGPAKAALPSILLSVINIC AHVLDDVAFGPAKAALPSILLSVINIC AHVLDVAFGPAKAALPSILLSVINIC AASGAVILLVAGGLALLVARINIC AASGAVILLVAGGLALLVARINIC AASGAVILLVAGLALLVARINIC AASGAVILLVAGLALLVARINIC AASGAVILLVAGLALLVARIPORSLYP AASGAVILLVAGLALLVARIPORSLYP AASGAVILLVAGGLALVARIPACHAALQV ALVVSKLPSHALLPPREKEDDVAFVAHALQV ANGNANCLEGEBODALA AARGAVILAVAGALGSERENDFRALQOY AVVSKLDEFGRAALQV AVVSKLDEFGRAALQV AVVSKLDEFGRAALQV AVVSKLDEFGRAALQV AVVSKLDEFGRAALQV AVVSKLDEFGRAALQV AVVSKLDEFGRAALQV AVVSKLDEFGRAALQV AVVSKLDEFGRAALGQLAVCHARINIC AARGAVALSVALGHRUNGAALQAVALGHRUNGUNGAALQARAALQV AVVSKLDEFGRAALQAVALGHLARLSVALGARAAL	ĺ .						
LETLAEIDFRLVSFLEAKAEMLIRGAHHYTCI LKLQERVLINVVIHLGDEDPRVRHVAAASL IRLYPKLFYKCDQGADPVVAVARDQSSVYL KLLMETOPPSHERVSTITRIYRGYNLLPSITD VTMENNLSRVIAAVSHELITSTTRALTFGCCE ALCLLSTAFPVCINSI.GWHGCYPPLASADESR KSCTVGMATMILTLLSSAWFLDLSAHQDAL LLAGNALAASAPKSLRSSWASEEBANPAATK QEEVWPALGDRALVFMVSQLFSSLLKVINIC AHVLDDVAPGPAIRAALPSLTNPFSLSPIRKR GEKERDGGASVFYSKKOSBASAASRQSDTS GFVTTSKSSILGSTYHLFSYLKHDVLSATHAL NYKVYLDLQNSTERKGGGRASALDVLSQILEL ATLQDIGKCVEELGYLKSCPSREPMMATVV VQULKTLGTGHLASQFDGLSSINPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNINVQAGQENDSTSGWFDVLQKVSTQLKT NLTSVTKORADKNAHNHIRLFEFLVIKALKQ YTTTTCVQLQKQVLDLAQLVQVATYQCLL DSDQVFIGFVLKQFEYIEVGGFRESEAIIPNIFF FLVLLSYSEYHSKGIGFRIGLQCHWSYCLL DSDQVFIGFVLKQFEYIEVGGFRESEAIIPNIFF FLVLLSYSEYHSKGIGFRIGLQCHWSYCLL DSDQVFIGFVLKQFEYIEVGGFRESEAIIPNIFF FLVLLSYSEYHSKGIGFRIGLQCHWSYCLL DSDQVFIGFVLKQFEYIEVGGFRESEAIIPNIFF FLVLLSYSEYHSKGIGFRIGLQCHWSYCLL DSDQVFIGFVLKQFEYIEVGGFRESEAIIPNIFF FLVLLSYSEYHSKGIGFRIGLQCHWSYCLL DSDQVFIGFVLKQFEYIEVGGFRESEAIIPNIFF FLVLLSYSERYHSKGIGFRIGLQCHWSYCLL DSDQVFIGFVLKQFEYIEVGGFRESEAIIPNIFF FLVLLSYSERYHSKGIGFRIGLQCHWSYCLL DSDQVFIGFVLKQFEYIEVGGFRESEAIIPNIFF FLVLLSYSERYHSKGIGFRIGLQCHWSYCLL DSDQVFIGFVLKQFEYIEVGGFRESEAIIPNIFF FLVLLSYSERYHSKGIGFRIGLQCHWSYCLL DSDQVFIGFVLKQFEYIEVGGFRESEAIIPNIFF FLVLLSYSERYHSKGIGFRIGLCCHWSYCLL DSDQVFIGFVLKQFEYIEVGGFRESEAIIPNIFF FLVLLSYSERYHSKGIGFRIGLCCHWSYCLL DSDQVFIGFVLKQFEYIERFILLYQYYCULL DSDQVFIGFVLKQFTISTLARDALDRODSTSTLE BSGKQKNLPEETFSRFLLQUVILLENGLUNGSTED VLSSIGLAGUNGSTSTLE BHARCHARLSTYNGCHSSFYNGLARAWAL SCHOOLH SKRIFT GRANDAL GARRANGERINGSTSTLE HISBGKQKNLPEETFSRFLLQVGILLEDIVT KQLKVEMSEQQHTFYCQLGTLLMCLHIFKS GMFRITAATATLRSDGGGGSFYALFRIN AASGLFIQAIQSRCENLSTTMLKKTLQCLEGG HLSQSGAVLTLYVDRLCTPFRVATARAMYDLL ACRAVEMLLAANLQSSMAQLPMEELNRIGEF LLSGGAVLTLYVDRLCTPFRVATAGAVLTELL SWHLLHEQPILSLDLQAGLOCCLALQLPGL WSVNSSTEPVIBLAGELVNSTTRYPDLWKLIG WSVNSSTEPVIBLAGELVNSTTRYPDLWATLE LSWHLHEQPILSUDLQAGLOCCLALQLPGL WSVNSSTEPVIBLAGELVNSTTRYPDLWATLE LSWHLHEQPILSUDLQAGLOCCLALQLPGL WSVNSSTEPVIBLAGELVNSTTRYPDLWWLG GESPFEEDTWATLGVVAQCPRINKLALGKRNSGVPA FLITLLRNIBILSARLPLVNSTTRYP							
ILKLQERVLINVVIHLLOEDEPRVRIPVARQOSSVYL KILIMETOPPSIFSVSTITRYRGYNLLPSID VTMENNLSRVIAAVSHEITSTTRALTFOCCE ALCILSTAFPVCIVSIG WHGVPPLSASDESR KSCTVGMATMILTILSSAWFPLDLSAHQDAT URAGNILAASAPKSIRSSWASEELARPAATK QEEVWPALGDRALVPMVSQLFSHLLKVINIC AHVLDDVAPGPARAALPSLTNPSISJPIRK GKEKEPGGQASVPLSKKGSEASAASRQSDTS GPVTTSKSSSLGSFYHLPSYLLHDVLKATHA NYKVTLDLQNSTEKFGGFRSALDVLSQILEL ATLQDIGKCVEELIG VYLKSPSEREPMATVC VQQLLKTLFOTNLASQPDGLSSNPSKSQGA QRLGSSSVPRGLYHCYMLLHFTQALADA SLRNNVQAGGENDTSGWFDVLQKVSTQLKT NLTSYTCNRAQNAHHRHIRLEFFLVIKALKQ YTTTTCVQLGKQVLDLAQLVQLRVNYCLL DSDQVFIGFVLKGFTEFFSFFLVIKALKQ YTTTTCVQLGKQVLDLAQLVQLRVNYCLL DSDQVFIGFVLKGFTEFFSFFLVIKALKG YTTTTCVQLGKQVLDLAQLVQLRVNYCLL DSDQVFIGFVLKGFTEFFSFFLVIKALKG CHKENEDK WKRLSRQLADILPMLAKQOMHH DSHEALAGVINTIFELARSSLRVDMLRSMF VTPNTMASVSTVQLWISGILALRVLISQSTED VISRGIGLSFSPYLISCTVINRLADGSTSTLE HSBGGGKRUPEETTSKFLLQLVGILLEIPKS GMFRITAAATRLFRSDGCGSFYTLDSLINK ARSMTTHPALVLLWCQLLLVNHTDYRW ARVQOTFKRHSLSTKLLSPQMSGEEEDSLAL AKLGMCNEFFFREGALLECVOUNTDLACL AKLGMCNEFFFREGALLECVOUNTDLACL AKLGMCNEFFFREGALLECVOUNTDLACL AKLGMCNEFFFREGALLSTCHURCHDLAC AKRUEMLAANLQSSMAQLPMEELNRIGEPS UKSRGJGRAGHLARVLINGSTSTLE HSBGGAVITLYVORLLCTFFRLACHLOLDLAC AKLGMCNEFFFREGALLSTCHURCHDLAC AKLGMCNEFFFREGALLSTCHURCHDLAC AKLGMCNEFFFREGALLSTCHURCHDLAC AKLGMCNEFFFREGALLSTCHURCHDLAC AKLGMCNEFFFREGALLSTCHURCHDLAC AKLGMCNEFFFREGALLSTCHURCHDLAC AKLGMCNEFFFREGALLSTCHURCHDLAC AKLGMCNEFFFREGALLSTCHURCHDLAC AKRUEMLAANLQSSMAQLPMEELNRIGEPS PPVSSHPLDGDGHVSLETVSPRLDWYPHLVK SQCWTRSDSALLEGAELVNRIPAEDMAPM MSEPNLSLLAFCLSLGMEESGGGKSALFEA AREVTLARNSGTVQCLAVHHUPGELPAEP AAYWSKLNDLFGDAALYQSLFTLARALQQY LVVSKLPSHLHPFREEDISTMQDSLSPS PPVSSHPLDGDGHVSLETVSPRLDWYPHLVK SQCWTRSDSALLEGAELVNRIPAELAVAVQDG EGLLSPERETVRFRASISEEEEDPDTTOPNEKYI TAACEMVAELAVSGTVQCLAVHHUPGELPAEP AAYWSKLNDLFGDAALYQSLFTLARALQQY LVVSKLPSHLHPFREEDISTKNQDSLSPS PPVSSHPLDGDGHVSLETVSPRLDWYPHLVK SQCWTRSDSALLEGAELVNRIPTERAVAVQDG EGLLSPERETVITFACASLTVCVINITERAVATURG GESPFEEDTERVTINVAXQGATSLVLSPLAG SWHLEBGPGGGGGGGGGGGALFEA AGEVTLARNSGTVQCEGRGAALTERA					:		
IRLYPKLFYKCDQGQADPYWAVARDQSSYVL KILIMETOPPSHEVSTITRTNYRGYLLPSITD V7MENNISRVIAANYSHELITSTTRALTIFGCCE ALCLLSTAPPYCURSI.GWHCGPYPLASADESR KSCTVGMATMIT.TLLSSAWPLDLSAHQDAL LAGORLAASAPKSLRSSWASEEBEAPPAATK QEEVWPALGDRALVPMYEQLFSHLLKVINIC AHVLDDVAPGARAALPSLTMPPSLSPIRK GKERPGGASVYJESYKKGGBASAASRQSDTS GPVTTISKSSLGSTYBLFSYLKLHDVLKATHAL NYKVTLDLQNSTERFGGFRSALDVLSQILEL ATLQDIGKCVEELIGYLKSCPSKEPMATVV VQULKTLGTTILAGOPGGLSSNDYSQUKSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKVSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKVSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKVSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKVSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKVSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKVSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKVSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKVSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKVSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKVSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKSTQLKT NLTSVTKORADKNAHMIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLAQLVQKSTQCLLALLX VXXILER LESDQVFIGFVLKQFEYTSCQFRESEAINPIFF FLVLLSYSTRENDGGRISCHVYCQLLLAQLX CKKERPOKWKLSRQLADLPLACLT BEQUE CKKERPOKWKLSRQLADLPLACLT COLKER LEVENSTANLSPOKOSTC LESDQVFIGFVLKQFETSRFLLQLVQHLLAQLX VTPNTMASVSTVQLUSGGILALLXALLQSTED VLSGGLAARLQLLLQTHVCQLULLANNTDYRWW ARVQCTKRHSISSTRLSTCYNNLTRYBLONDLX ARSMITHFALVLLWCQLLLVNNTDYRWW ARVQCTKRHSISSTRLSSPOKOSTNTDENSTLL EHSBGKQKNLPEETFSRFLLQLVGHLLKNTTOVAL ARSMITHFALVLLWCQLLLVNNTDYRWW ARVGCTKRHSISSTRLSSPOKOSTNTDENSTLL CHIRCTORADKNAMMIRLFQFTVARAMVDLL ACRRVEMLLAANLQSSMAQLPMEELNRNGVSP LYVSKSHPLDGGGGGSALTSTYMAKATQLARAMVDLL ACRRVEMLLAANLQSSMAQLPMEELNRNGVSP LQSGGAVLTTVYDRLCTPFRVATLGAUNTALAQU NSVKSTEPVIBLORGCCCLALQLPGL WSVVSSTEPVIBLAGELVONTYRYPGLAMMIRLY AASGATALAQUSRCENASTHTNKALSEEBEEVDPHTQNNKYT TAACEMYABACLQCQPRIKKLALALDTRFGGK LSWRLBGDFGGAALYGARANATHLUTYQAWD PVSLSPATTOALSHEKLLLQQDPPGLYKLELGGM							
KLLMHETQPPSHFSVSTITRIVTRGYNLLPSTID VTMENNLSRVLAAVSIELITSTTRALTFGCCE ALCLISTAFPVCIWSLGWHCGYPPLSASDESR KSCTVGMATMILTLISSAWHCGYPPLSASDESR KSCTVGMATMILTLISSAWHCGYPPLSASDESR KSCTVGMATMILTLISSAWHCGYPPLSASDESR KSCTVGMATMILTLISSAWHCGYPPLSASDESR KSCTVGMATMILTLISSAWHCGYPPLSASDESR KSCTVGMATMILTLISSAWHCGYPPLSASDESR KSCTVGMATMILTLISSAWHCGYPPLSASDESR GPEVTSKSSLGALAYMCGAWHLKATHA OPEN AND ALL AN	1				İ		
VTMENNLSRVIAAVSIELITSTTRALTGCCE ALCLISTAFPCUNSLGWIGGNEGLASDESS KSCTVGMATMILTLSSAWFPLDLSAHQDAL ILAGNILAASAPKSLRSSWASEANPAATK QEEVWPALGDRALVPMVEQLFSHLLKVINIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRK GKEKEPGQJASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSYYHLPSYLKHDVLKATHA NYXVTLDLQNSTEKEFGGFLRSALDVLSQILEL ATLQDIGKCVEELIGYLKSCFSREPMMATVC VQQLLKTLFGTINLASGPGDLSSPSSQGRA QRLGSSSVRGLYHYCFMAPYTHFTQALADA SLRINMYQAEDGENDTSGWBVDLGVSTQLKT NLTSYTKNRADKNAIFHNHRLFSPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYEVQQPRESEAIIPNIFF FLVLLSVERYHSKQIGHFKIQLCDGIMASGR KAVTHAPALQPVIBLIFLYGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFLVQQ CIKENDENWKRLSRQIADIHLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAULISQSTED IVLSRQGELSFSPYLISCTVINRLRDGDSTSTILE EHSEGKQIKNLPEFTSFRFLQUILEDIDTT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRITAAATRLFRSDCGGGSFT/LDSLILR ARSMITHPALVLUCQUILLIVHTDYRWW AEVQQTFKRHSLSSTKLLSQMSGEEDSDLA AKLGMCNBETVRRGALLEFCDYVCQNLHDSE HLTWLLVNHQDLISLSHEPVVQDFLONSLRS AASGLFQAIGSGRCENLSTPHKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDLL ACREVEMLLAANLQSSMAQLPMELNRIQES LOSSGLAQRHQRIVSTLDEFRELSTMQDSLSPS PFVSSHPLOGDGHVSLETTVSRRCLAUCHYVCHLHOS SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEPNISLLAPCLSGMSEESGGKSALFEA AREVTLARVSGTVQLPAVHLVKY SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEPNISLLAPCLSGMSEESGGKSALFEA AREVTLARVSGTVQLPAVHLYGPELPAEP AAYWSKLNDLFGDAALVQSLFPLARALAQY LVVVSKLPSFHLHPPEKKEDVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQPGL WSVVSSTEFVTHACSLIVCVFHLEAVAVQPG EQLLSPERRTNTPKAISEEGEVDPNTQPKYI TAACEAVAEMBWSSLOSJLALGHERNSGVPA AFITPLERMISLARLLYNKYTRVPPLVWKLIG WSPKPGGDFGTAFFEIPVEFLIGEKEVKEFTYR INTIGWTSRTQFEETWAILLGVLVTTQPLVME QEESPPEEDTERTQINALAVAQFG EQLLSPERTTNTPKAISEEGEVDPTTQPKKIT TAACEAVAEMBWSSLOSJLALGHERNSGVPA FITPLERMISLARLPLYKKATTRACHAVAQPG EQLLSPERTTNTPKAISEEGEVDPTTQPKKIT TAACEAVAEMBWGSLOSJLALGHERNSGVPA FITPLERMISLARLPLYKKATTRACHAVAQPG EQLLSPERTTDTPKAISEECETVPTTQPKKIG SAMT	1	[KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD
KSCTVGMATMILTILSSAWFFLDLSAHQDAL ILAGNILAASPKSLRSSWASPKSLNPSAKSLYNIC AHVLDDVAPGPAIKAALPSLTNPPSLSPIRK GKEKEPGPQASVPLSPKKOSEASAASRQSDTS GPVTTSKSSSLOSYPHLPSYLKHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDICKCVEELGYLKSCFSREPMMATVC VQULKTLFGTINLASQPGDLSYPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHETQALADA SLRNMVQAEDENDTSGWFDVLQKVSTQLKT NLTSYTKNRADKNAISHHRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQRYNYCLL DSDQVFGFVLKQFEYEVGQFRESEAIPNIFF FLVLLSYERYHSKQIOPKIQLCDGMASGR KAVTHAPALQPIVIDLFVLRGTNKADAGSGR KAVTHAPALQPVVDLLTLAFTNKADAGSGR KAVTHAPALQPVVDLLTLAFSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALKQQMSTLLSGNF EHSGKQIKNLPETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHHFKS GMFRRTAAATRLFRSDGGGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVMHTDYRWW AEVQQTFKRHSLSSTKLLSQMSGEEDSDLA AKLGMCNEETVRGALLFCDYVCQNLHDSE HLTWLLVNHQDLISLSHEPVVQFISAVHNS AASGLFQAIGSGCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDLL ACRRYEMLLANLQSSNAGLPELDSLSPS PPVSSHPLGGGMYSETVSPREDNWYVHLVK SQCWTRSDSALLEGAELVRRPAEDMNAFM MNSEFNISLAPCLSIGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHYGPELPAEP AAYWSKLNDLFGDAALVQSLFPLARALAQY LVVVKKPSHLHHPPEKEKDVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQPGL WSVVSSTEFVTHACSLIVCVFFLEAVAVQDG GQLJSPERRTNTPKAISEEGEVDPNTONPKYI TAACEAVABEMVSSLOSJALGHENSIGVPKY ITALGENSTERVESUS CALLGHENSIGVPK ITTLLENNISLAPLLSLAPLLSHELDVFLYDVKLG WSVVSSTEFVTHACSLIVCVFFLEAVAVQDG GQLJSPERRTNTPKAISEEGEVDPNTONPKYI TAACEAVAGEMVSSLOSJALGHENSIGVPK ITTLLSHMISLAPLLSLAPLLSHALAQY LVVVKKPSHLHHPPEKEKDVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQPRKYNGY TAACEAVAGEMVSSLOSJALGHENSIGVPK ATSIGNIFGGEGGAMYSKENIATHHLYQOWD PVPSLSPATTGALISHEKLLLQURKALDTFGKK LSIRGIVEQEIQAMYSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQURPRELGSMS							
ILAGNILAASAPKSLRSSWASEEANPATK QEEVWPALGDRALVPMOQLISHLLKVINIC AHVLIDDVAPQPAKAALPSI.TNPPSLSPIRKK GKEKEPGQASVPLSPKKGSEASAASRQSDTS GPVTTSKSSSLGSFYHLESYLKLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATI_QDIGKCVEELIGYLKSCSFREPMMATVC VQULLKTLFGTNLASQFPGLSSNPSKSGGRA QRIGSSSVPGCYHLYCMPAYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSYKNADAKNAHNHEPEPLYIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFYLKQFEYTEVGGFRESAHIPNIFF FLVLLSYRSYHSKQIGIPKIQLCDGIMASGR KAVTHAPALQPIVHDLFVLRGTNKADAGKE LETQKEVVYSMLTLIQYHQVLEMFILVLQQ CHEENEDKWRLSRQADDILPMLAKQOMHI DSHEALGVLNTLFELLASPLYDMLLRSMF VTPNTMASVSTVQLWISGILALBVLISQSTED IVLSRIGELSFSPYLLISCTVHLRGDGSTSTLE EHSEGKQIKNLPETTSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCHHFKS GMFRRITAAATRLFRSDGGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQOTPKRHSLSSTKLLSGVGGEESDLA AKLGMCNRETVRRGALILFCDYVCQHLHDSE HLTWLINVHIODLISLSHEPPVQDFISA VHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTI, VYDRLLCTFREVLARMVDIL ACRRVEMILAANLQSSMAQLFMEELNRIGSY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGGHYSLLTVRPEKDMAPFM MNSEPNLSLLAPCLSLGMSEISGGGKSALFEA AREVILARVSGTVQQLPAVHHYFQPELPAEP AAYWSKLNDLFGDGHYSLETVRJENMAPFM MNSEPNLSLLAPCLSLGMSEISGGGKSALFEA AREVITLARVSGTVQQLPAVHHYFQPELPAEP AAYWSKLNDLFGDGHYSLETVCHPILEAVAVQPG EQLLSPERTNTFKAISEEEEEVDPTNQPKYI TAACEMVAEMYSLLOPAGLUPCLU WSVLSSTEFFYHACSLTLYCHFILEAVAVQPG EQLLSPERTNTFKAISEEEEEDTRYPRYNAPHY TAACEMVAEMYSLLOPAGLUPCLUPCLU WSVLSTEFFYHACSLTLYCHFILEAVAVQPG EQLLSPERTNTFKAISEEEEEVDPTNQPKYI TAACEMVAEMYSLLOPAGLUPCLUPCLUPCLUPCLUPCLUPCLUPCLUPCLUPCLUPC							
QEEVWPALGDRALVPMCQLESHLLKVINIC AHVLDDWAPGPAKAAJPSI.TNPSLSPIRRK GKEKEPGGQASVPI.SPKKGSEASAASRQSDTS GPVTTSKSSSLGSFYHLPSYLKLEHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQUEL ATLQDIGKCVEELI.GYLKSCFSREPMATVC VQQLLKTLGTINLASQPFGLSSNPSKSQGRA QRLGSSSVRPGLYHYCFMAPYTHFTQALADA SLRNMYQAGEONDTSGWPU.KQVSTQLKT NLTSYTKNRADKNAHNHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFYLKQFEYTWEGGPRESEAIIPNIFF FVLLSYERYHSKQIGFPKIQCDOMASGR KAVTHAIPALQPIVHDLFYUKQTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFU.VQQ CHENEDKWKRLSRQLDIDLFMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGIALBLVILSQSTED IVLSRQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQPGLGTLLMCLHIFKS GMFRRIAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWQCILLUMYTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREFVRRGALLHCDVYCQHLDBSE HLTWLINVHIDQDLISSJEPPYQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTT.YVDRLLCTRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRIQRLSYSLTYVDRLATQCLEGI HLSQSGAVLTT.YVDRLLCTRVLASMVDUL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRIQRLSYSLDRANSTMODSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRFAEDMNAFM MNSEPNISLLAPCLSJCGMESIGGQKSALFEA AREVTLARVSGTVQQLPAVHHYCPGELPAEP AAYWSKLNDLFGDAALYGELPLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVAATLAAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEPVTHACSLITYCHPLELAVQNG FUTLERNIBISLARLPLNSSTRYPRYPONFYY TAACEMVAEMVSSLOSJALGHKRNSGYPA FUTLLRNIBISLARLPLNSSTRYPLVWLIG WSVPSGEGGTAFPEPPYGPFLQEKEVFKEFIYR INTLGWTSRTOFEETWATLLGALNGTPLWFIT ACACEMVAEMVSSLOSJALGHKRNSGYPA FUTLLRNIBISLARLPLNSSTRYPLVWLIG WSVPSGGGGTAFPEPPTFLQEKEVKEFIYR INTLGWTSRTOFEETWATLLGALDTFTGGKK LSIRGIFVSGLOAMVSKRENNATHHLYQAWD PVPSLSPATTGALSLIHCALLLQNPERELGSMS	1						
AHVLDDVAPGPAIKAALPSI.TNPPSLSPIRRK GKEKEPGGASVPLSPKEASAASROSDTS GPVTTSKSSSLGSFYHLPSYLKLHDVI.KATHA NYKVITDL(NSTEREGGILRSALDVI.SQILEL ATLQDIGKCVEELIGYLKSCFSREPMMATVC VQULIKTLFGTHLASQFDGLSSNFSKSGRA QRIGSSSVPROLYHYCTMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSYTKINADKNAHMIEPEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYYEVQGPRESEAILPNIFF FLVLLSYBRYHSKQIGFBKIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVYSMLTLIQHYQVLEMFILVLQQ CHKENEDKWKRLSRQIADDILPMLAKQOMHI DSHEALGVINTLFELLASPRYDMLLRSMF VTPNTMASVSTVQLWISGILAILEVILGSTED IVLSSIQGLSFSPYLISGVTURLRDDGSTSTLE EHSEGKQIKNLPEETFSRFLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLMCLHIFKS GMFRRTIAAATRLFRSDGGSSFYTLDSLNLR ARSMITHPALVLLWCQILLLUNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQHLHDSE HLTWLINNIHODDLISSLEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFREVLARMVDIL ACRRVEMLLAANLQSSMAQLFMEELNRIQES LQSSGLAQRHQRLYSLDRFRLSTMQDSLSPS PPVSSHPLJOGGHSVLTVFLVRXVLLCTFREVLARMVDIL ACRRVEMLLAANLQSSMAQLFMEELNRIQES LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLJOGGHSVLSTVFLARMLAQY LVVVSKLPSHLHPPEKERGINGENGGGSGYALFAA AREVITLARVSGTVQQLPAVHHYFOPELPAEP AAYWSKLNDLFGDGALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKERDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFFVTHACSLTLYCHFLLEAVAVQPG EQLLSPERTINTFKAISEEEEEDPHTYQNFKYI TAACEMVYBAMVSLLGYKRPLLORKPRYPA							
GKEKEPGEQASVPLSPKKGSEASAASRQSDTS GPVTTSKSSLGSYPLSYLLHDVLKATHA NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL ATLQDIGKCVEELIGYLKSCFSKEPMMATVC VQLLKITLFOTNLASQPGCLSSNPSKSQGRA QRLGSSSVRFGLYHYCFMAFYTHFTQALADA SLRNWVQAEQENDTSGWFDVLQKSTQLKT NLTSVTKNRADKNAIFNHIRLFEPLVIKALKQ YTTTTCVOLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFFYEVGGFRESEAHPNIFF FLV1LSYERYHSKQHIGHKIQLCGIMASGR KAVTHAPALQPYHEULFVLRGTKKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFLVLQQ CHKENEDKWKRLSRQIDLIPMLAKQOMHI DSHEALGVLNTLFEILAPSLRFVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIGELSFSPYLISCTVINNLRDGDSTSTLE EHSEGKQIKNLPEFTFSRFLLQLVGLLEDIVT KQLKVEMSEQQHTFYCGLGTLLMCLHHFKS GMFRITAAATRLFRSDGGGGSTYLDSLNLR ARSMITTHFALVLLWCGLLLVNHTDYRWW AEVQQTFKRHSLSSTKLLSPOMSGEEDSDLA AKLGMCNREIVRGALLFCDYVCQNLHDSE HLTVLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFQAIGSRCENLSTIFMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFRFVLARWYDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAGRIGNLYSLDERISTMODSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLTCPFRVLARWYDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAGRIGNLYSLDERISTMODSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGELVNRIPAEDMNAPM MNSEFNISLLAPCLSLGMSEISGGKSALFEA AREVITLARVSGTVQQLPAVHHVOPELPAEP AAYWSKINDLFGDAALYQSLFTLARALAQY LVVVSKLPSHLH.PPEKEKDIVKFVVATLEAL SWHLIBEQPIS.DLQAGLDCCCLALQLPGI WSVVSSTEFVTHACSLLYCHFILEAVAVQPG EQLLSPERTNITYKAISEEEEEVDPNTQNPKYI TAAACEMVAEMVESLQOPRINKTLALOTYRFOKK WSVSSTEFVTHACSLLYCHFILEAVAVQPG EQLLSPERTNITYKAISEEEEEVDPNTQNPKYI TAAACEMVAEMVESLQOPRINKTLALOTYRFOKK NTILGWTSRTQFFETWATLLGULTVTOPLVME QESSPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPROKPLKALADTRFGRK LSIRGIVSQEQAMYSKRENIATHHLYQAWD PVPSLSPATTGALLSIEKELLLQNRPERELGSMS	1						1 3
GPVTTSKSSILGSFYHLPSYLKLHDVLKATHA NYKVTLDLONSTEKFGGFLRSALDVLSQIE L ATLQDIGKCVEEILGYLKSCFSREPMMATVC VQQLLKTLFGTNILASQFDGLSSNPSSQGRA QRLGSSSVPGLYHYCPMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAIPHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGGFRESEAIIPHFF FLVLLSYERYHSKQIGIPKIIQLCDGIMASGR KAVTHAPPALQPVIPDLFVLRGTTSKADAGKE LETYKEVVSMLLRLIQWQLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHBALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILALRVLISQSTED IVLSRQGLSFSFYLISCTVINBLRDGDSTSTLE EHSEGKQUKNLPEETFSFFLLQVGLLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRRIFAAATRLFSSFLLQVGLLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRRIFAAATRLFSSFLLQVGGLLEDIVT KQLKVEMSEQQHTFYCQBLGTLMCLHIFKS GMFRRIFAAATRLFSSFLLQVGGLLEDIVT KQLKVEMSEQGHTFYCQBLGTLMCLHIFKS GMFRRIFAAATRLFSSFLLQVGGLLEDIVT KQLKVEMSEQGHTFYCQBLGTLMCLHIFKS GMFRRIFAAATRLFSSFLLQVGGLEDIVA ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTFKRISLSSTKLLSPQMSGEEDSDAL AKLGMCNEFVRRGALIFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFQAIGNSCENSTPMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLAAANLQSSMAQLPMEEINRIGEY LQSSGLAGRRGRLYSLFRIFMGMSTLSFS FPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCVTTRSDSALLEGBAUKNIPAEDMNAFM MNSEPNLSLAPCLSI GMSEISGGQKSALFFA AREVTLARVSGTVQQLPAVHHYOPELPAEP AAYWSKLNDLEGDAALVQSLFTLARALAQY LVVVSKLPSHLHLPFEKEKDIVKFVVATLEAL SWHLHEGPUPSLDLQAGLOCCLALQLPGI WSVVSSTEFVTHACSLIVCVHFILEAVAVQPG EQLLSPERTNTFKAISEEEEVDPNTQNPKYI TAAACEMVAEMVESLGSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGGFGTAPPPETQEKEVFKEFTYR NTLGWTSRTOFFETWATLLGVLVTOPLVME QEESPPEEDTERTQNVAXQATISLVLSAMT VPVAGNPAVSCLBQQPRNRPLKALDTRFGRK LSIRGIVSQEQAMVSKRENIATHHLYQAWD PVPSLSPATTGALLSIEVELJLQNFRELGSMS							
NYKVTLDLONSTEKFGGFIRSALDVLSQUEL ATLODISC VEEL GYLKSCPSREPMMATVC VQQLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRFGLYHYCFMAPYTHFTQALADA SLRNMVQAQEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAIFNHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFFYEVGQFRESEAIIPNIFF FLVLLSYERYHSKQIIGIPKIQLCGIMASGR KAVTHAPALQPYHEDLFVLRGTNKADAGKE LETQKEVVVSMLIRLIQYHQVLEMFILVLQQ CHKENEDK WKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAFSLRFVDMLLRSMF VTPNTMASVSTVQLWISGILALIRVLISQSTED IVLSRQGESSSPYLISCTVINILRDGDSTSTILE EHSBGKQIKNLPEFTSRFILQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTILMCLHIFKS GMFRITAAATRLFRSDGGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVINITDYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRGALILFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPVQDFISAVHRNS AASGLFQAJQSRCENLSTITMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFPRVLARMVDIL ACRRVEMLLANILGSSMAQLPMEEINRIQEY LQSSGLAQRIQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEPNLSLAPCLS GMSEISGQGSASPTEA AREVITLARVSGTVQQLPAVHHVFQPELPAP AAYWSKLNDLFGDAALYQSLFTLARALAQY LVVVSKLPSHLHLPEKKENDVKFVVATLEAL SWHLHEQPPLSLDLQAGLDCCCLAQLPGL WSVVSSTEFVTHACSLTVCVHFILAVALQPG GQLISPERTNTPKAISEEEEEVDPNTQPKYN TAACEMVAEMVESLQSVLALGHKRNSGVPA FILTPLIRNIISLARLPLVNSYTRVPLVWKLG WSPKPGGDGFGTAFPEPVEFIQEKEVFKEFTYR NTILGWTSRTQFEETWATLLGVTVPLVWKLG WSPKPGGDGFGTAFPEPVEFIQEKEVFKEFTYR NTILGWTSRTQFEETWATLLGVTVPLVWKLG WSPKPGGDGFGTAFPEPVEFIQEKEVFKEFTYR NTILGWTSRTQFEETWATLLGVAQHTSLVJSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGK LSIRGIVSQEQEAMVSKRENIATHLYQAWD PVYSSLPATTGALSIASALTLANDRYPLLVSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGK LSIRGIVSQEQEAMVSKRENIATHLYQAWD PVYSSLPATTGALISTEKLLLQNPERELGSMS							
ATLODIGKCVEELIGYLKSCFSREPMMATVC VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVRPGLYHYCPMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAIHNHIRLFEPLVKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFFEVEVGGFRESRAIPNIFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTTNKADAGKE LETQKEVVSMLLRLQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHBALGVINTLFEIL APSLRPVDMLLSSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED INLSRIQELSFSPYLISCTVINRLABGDSTSTILE EHSEGKQIKNLPEETFSFFLLQLVGLLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRRITAAATRLFSSFLLQLVGLLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRRITAAATRLFSSFLLQLVGLLEDIVT KQLKVEMSEQQHTFYCQDLGSTLTMCLHIFKS GMFRRITAAATRLFSSFLLQLVGLLEDIVT KQLKVEMSEQGGSFYTLDSLNLR ARSMITTHAALVLLWCQULLJVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQNSGEEDSDLA AKLGMCNREIVRRGALLFCDYVCQNLHDSE HLTVLIVNHQDLISLSHEPPVQDFSAVHRNS AASGLFJQAIGSGFCENSITMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRRQRLYSLDRFRLSTMQDSLSPS PPVSSIPLDGDGHVSLETVSPDKDWYVHLVK SQCVTRSDSALLEGAELVNRIPAEDMNAFM MISSPNLSLAPCLSCMSESGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLFTLARALAQY LVVVSKLPSHLHLPFEKENDIVKFVVATLEAL SWHLHERDPLSDLJQAGLDCCCLALQLPGL WSVVSTEFVTHACSLIVCVHFILEAVAVQPG EQLLSPERTNTPK AISEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLAARLPL VNSYTRVPLVWKLG WSPKPGGDFGTAPPETQLEKEVFKEFTYR NTLGWTSRTOFFETWATLLGVLVTQPLVME QEESPPEEDTERTQNTVAQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKAUDTRFGRK LSIIRGIVSQEIQAMVSKRENIATHLYQAWD PVPSLSPATTGALLSIEKLLLQNRPERLEGSMS	1.						L
VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA QRLGSSSVPROLLYFMAPYTHFTQALADA SLRNMVQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAHINHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYEVGQFRESEAIIPNIFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWRLSRQIADILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQVGILLEDIVT KQLKVMSEQQHTPYCQELGTLLMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLWCQILLLVHTDYXWW AEVQQTPKHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLINNHIQDLISHPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRILCTFRVLARMYDIL ACRRVEMLAANLGSMAQLPMEEINRIQEY LQSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLOGDGHVSLLDTFRSDMYDVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVILARVSGTVQLAYHHVPGELPAEP AAYWSLINDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKKKDIVKFVVATLEAL SWHLHEQIPLSLDLQGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLSPBFRTNTFRASEEEEEVDPTONYPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLITPLLRNIISLARLPLNSTYTRYPPLVWKLG WSVRSGGFGTAFEPIEVEFLQEKEVKFEFTYR INTLGWTSRTQFEETWATLLGVLYTQPLVME QESPPPEDTERTQNLAVAQATISLYLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPALSPATTGALISTATHLYQAWD PVPALSPATTGALISTATHLIYQAWD	İ	İ					
ORLGSSSVRÖLYHYCFMAPYTHFTQALADA SLRNMYQAEGENDTSGWFDU,QKVSTQLKT NLTSVTKNRADKNAIHNHIRLFEPLVIKALKQ YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVPIGFVLKQFEVGYFRESEAIPNIFF FLVLLSYERYHSKQDIGPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFLAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLDGDSTSTILE EHSEGKQKNLPEETTSRFLLQLVGGILEDIVT KQLKVEMSEQQHTFYCOELGTLLMCLHHFKS GMFRRITAAATRLFRSDGCGSFYTLDSLNLR ARSMITHPALVLLW CQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSFOMSGEEDSDLA AKLGMCNEFVRRGALILFCDYVCONLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACRRVEMLAANLQSSMAQLPMEELNRLQEY LQSSGLAGGAVLTLYVDRLLCTPFRVLARMYDIL ACRRVEMLSAANLQSSMAQLPMEELNRLQEY LQSSGLAGHQRLYSLLDRFRLSTMQDSLSFS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEPNISLLAPCLS GMSEISGGGKSALFEA AREVTLARVSGTVQQLPAVHHYQPELPAEP AAYWSKINDLFGDAALYQSLTLARALAQY LVVVSKLPSHLHIPPEKEKDIVKFVVATLEAL SWHLHBOPLSLDLQAGLDCCCLALQLPGL WSVVSSTEPVHACSLIFVATRALAAQY LVVSKLPSHLHIPPEKEKDIVKFVVATLEAL SWHLHBOPLSLDLQAGLDCCCLALQLPGL WSVVSSTEPVHACSLIFVATRASLECEEVPDFINTONPKYI TAACEMVAEMVESLQSVLALGHKRNSGYPA FLITELRMIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFFEIPVEFLQEKEVFKEFTYR NTLGWTSRTQFEETWATLLGVLYTQPLVME QESPPFEDTERTQINLAVAVAATISLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALLSVLLLQNPREELGSMS	ĺ						VOOLLKTLEGTNLASOFDGLSSNPSKSOGRA
SLRNMYQAEQENDTSGWFDVLQKVSTQLKT NLTSVTKNRADKNAIHNHIRLFEPLVIKALKQ VTTTTCVQLQKQVJLDLLAQLVQLRYNYCLL DSDQVFIGFVLKQFEYEVGQFRESEAIIPNIFF FLVLLSVERYHSKQIPKIQLCDGIMASGR KAVTHAPALOPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQLEMFLUQQ CHKENEDKWKKLSRQIADILIPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLSGIALIRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETTSRFLLQVGILLEDIVT KQLKVEMSEQQHTPYQCQELGTLLMCLHIFKS GMFRRTAAATRLFRSDGCGGFYTLDSLNLR ARSMITTHPALVLLWCGILLIVHTYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEDISDLA AKLGMCNREIVRRGALLIFCDYVCQNLHIDSE HLTWLINVHIQDLISHPPYQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRILCTPFRVLARMYDIL ACRRYEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDFFRLSTMQDSLSFS PPVSSHPLDGGGHVSLETVSPDKDWYHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGG(RSALFEA AREVILANSGTVQLAVHHVPGELPAEP AAYWSKINDLFGDAALVGSLPTLARALAQY LVVSKLPSH.HLPDEKKEDIVKFVVATLEAL SWHLHEQIPLSLDLAGLDCCLALQLFGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLISPERRTNTFRAEEEEEVDPTINQPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLIPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFFEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFEETWATLLGVLYTQPLVME QESPPEEDTERTQINAVAVAQATISLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPALSPATTGALISLLLQNPRELGSMS							
NITSVTKNRADKNAHNHIRLFEPLVIKALKQ YTTTTCVQLQKUDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF FLVLLSYBRYHSKQIIGFRIQLCDGIMASGR KAVTHAPALQPVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADILPMLAKQQMHI DSHEALGVVINTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQL WISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYVQGLGTLLMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLALR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRRISLSSTKLLSPQMSGEEEDSDLA ARLGMCNREIVRRGALLIFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVOPFISAVHRNS AASGLFQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMYDIL ACRRYEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQVSLLDFRFLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEPNLSLLAPLSLGMSEISGG(KSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHPEPKEKDIVKFVVATLEAL SWHLHEQPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNITPKAISEEEEEVDPNTQNPFKYI TAACEMVAEMVESLQSVLALGFIKRNSGVPA FLTPLLRNIISLARLP VNSYTRYPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVKKLG WSPKPGGDFGTAFPEIPVEFLQEKEVKKLFTYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESSPEEDTERTONVLAVQAITSLVLSAMT VPVAGRPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMYSKENATHHLYQAWD PVPSLSPATTGALISHEKLLLQUNPERELGSMS							
DSDQVFIGFVLKQFEYEVGQFRESEAIPNIFF FLVLLSYERYHSKQIIGPKIIQLCDGIMASGR KAVTHAPALQPIVHDLFVLRGTINKADAGKE LETQKEVVVSMLIRLIQYHQVLEMFILVLQQ CHKENEDKWKLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPVLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHITFYCQELGTLLMCLHHIFKS GMFRRITAAATRLFRSDGGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDVVCQNLHIDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFQAIGSSCENLSTFINLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMQLPMEELNRIQEY LOSSGLAQRHQRLYSLDBFRISTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLACLSIGMSEIGGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQFELPAEP AAYWSKLNDLFFDAALVQSLPTLARALAQY LVVVSKLPSHLHPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFTHACSSITYCVHFILEAVAVQPG EQLLSPERRINTFKAISEEEEEVDPNTQNFKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLALPLYNSYTRYPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFEETWAILLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRIKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQQNISTLVLSAMT VPVAGNPAVSCLEQQPRIKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNIPERELGSMS							
FLVILSYERYHSKQIIGPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIOELSFSPYLISCTVINRLRDGDSTSTLE EHSGEKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLHIFKS GMFRRITAAATRLFRSDGGGGSFYTLDSLNLR ARSMITHPALVLWCQILLIVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHOJLISLSHEPPVODFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLICTPFRVLARMYDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLTVSPDKDWYVHLVK SQCWTRSDSALLEGABLVNRPAEDMINAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVITLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLTLARALAQY LVVVSKLPSHLHIPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLAIQLPGL WSVVSSTEFYTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRYPPLVWKLG WSPKPGGDFGTAFPEPVFFLQEKEVFKEFTYR INTLGWTSRTQFEETWATLLGULYUPLVME QEESPPEEDTERTQINVLAVQATISLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALJSHEKLLLQNPEPELGSMS							
KAVTHAPALOPIVIÐLEVLROTINKADAGKE LETQKEVVVSMILRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIAÐIILPMLAKQOMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAIRVLISQSTED IVLSRIGELSFYLLSCTVINRLRDGDSTSTILE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHITFYCQELGTILMCLHIFKS GMFRITAAATRLFRSDGGGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTFKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNRETVRRGALLIFCDVVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDFRRLSTMQDSLSPS PPVSSHPLDGGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLAPCLS.LGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTINTFKAISEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIBSLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVVFUPLVMEL QEESPPEEDTERQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
LETQKEVVVSMLIRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVINTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTILE EHSEGKQKINLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTPYCQELGTILMCLHIPKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDLL ACRNVEMLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEPNLSLLAPCLSLGMSEISGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDIQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIEGIVEQEIQAMVSKRENIATHHLYQAWD PVYSLSPATTGALISHEKKLLLQINPERELGSMS		ļ					
CHKENEDKWKRLSRQIADILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSQQHTFYCQELGTLLMCLHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQLLLVNHTDYRWW AEVQQTPKHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALIFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTFFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELWRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGGKSALFEA ARREVTLARVSGTVQQLPAVHHVPQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIFLSDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRINTPKAISEEEEEEVDPNTQNFKYI TAACEMVAEMVESLQSVLAIGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQAFFETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKKLLLQINPERELGSMS							
DSHEALGVLNTTLFEILÄPSSLRPVDMLLRSMF VTPNTMASVSTVQLWISGILAIRVLISQSTED IVLSRIGELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCOELGTLLMCLHIFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMTTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEPNLSLLAPCLSLGMSEISGGGKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHPPEKEDIVKFVVATILEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLYTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGRPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS	1	}					
VTPNTMASVSTVQLWISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTILE EHSEGKQKINLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRIFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALLIFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGABLVNRPAEDMNAFM MNSEPNLSLLAPCLSLGMSEISGGKSALFEA AREVTLARVSGTVQQLFAVHHVFQPELPAEP AAYWSKLNDLFGBAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQPPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEVDPNTQNFKYI TAACEMYAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGSPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPPERELGSMS							
IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEFTSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCOELGTLLMCLHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVITLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVSSLPSHLHLPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLTYCVHFILEAVAVQPG EQLLSPERTINTPKAISEEEEEVDPNTQNFKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS	1	·					
EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTILMCLHIJFKS GMFRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNRETVRGGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLDFFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MINSEFNLSLLAPCLSLGMSEISGG(KSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLITPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS							
KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRIFRSDGCGGSFYTLDSLNLR ARSMITHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSIPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEPNLSILAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKINDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							3
GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITHPALVILWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFJQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLTYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMYAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRINKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1						
AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMYDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNIPERELGSMS							
AKLGMCNREIVRRGALILFCDYVCQNLHDSE HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1			,			
AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1	'					AKLGMCNREIVRRGALILFCDYVCQNLHDSE
HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1	i					
ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS	1						
PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS						8	
SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS							PPVSSHPLDGDGHVSLETVSPDKDWYVHI VK
MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS					·		
AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS	1						MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA
AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS							AREVTLARVSGTVQQLPAVHHVFQPELPAEP
SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							AAYWSKLNDLFGDAALYQSLPTLARALAQY
WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1						SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL
TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG
FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							
WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1						
INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS		[
QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	1	·					
VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS							OEESPPEEDTERTOINVLAVOAITSI VI SAMT
LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQINPERELGSMS	·				* -		VPVAGNPAVSCLEQOPRNKPLKALDTRFGRK
PVPSLSPATTGALISHEKLLLQINPERELGSMS	1						LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD
YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE							PVPSLSPATTGALISHEKLLLQINPERELGSMS
	L	<u> </u>					YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of, peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL LELYSRWILPSSSARRTPAILISEVVRSLLVVS DLFTERNQFELMYVTLTELRRVHPSEDEILAQ YLVPATCKAAAVLGMDKAVAEPVSRLLESTL RSSHLPSRVGALHGILYVLECDLLDDTAKQLI PVISDYLLSNLKGIAHCVNIHSQQHVLVMCAT AFYLIENYPLDVGPEFSASIIQMCGVMLSGSE ESTPSIIYHCALRGLERLLLSEQLSRLDAESLV KLSVDRVNVHSPHRAMAALGLMLTCMYTG KEKVSPGRTSDPNPAAPDSESVIVAMERVSVL FDRIRKGFPCEARVVARILPQFLDDFFPPQDIM NKVIGEFLSNQQPYPQFMATVVYKVFQTLHS
428	1778	A	3449	3	430	TGQSSMVRDWVMLSLSNFTQRAPVAMATWS LSCFFVSASTSPWVAAILPHVISRMGKLEQVD VNLFCLVATDFYRHQIEEELDRRAFQSVLEV VAAPGSPYHRLLTCLRNVHKVTTC NSRPSPSAALVEVLLRSGSTFPHTVSGGWAA
				-		WGPWSSCSRDCELGFRVRKRTCTNPEPRNGG LPCVGDAAEYQDCNPQACPVRGAWSCWTS WSPCSASCGGGHYQRTRSCTSPAPSPGEDICL GLHTEEALCATQACPEGWS
429	1779	A	3464	583	3	DALDRRYLERCHPAAGGWVGEGE*ALCQKT/ RFSGVLEPPLPSLKDGGRFPAWT*RSCSKSLR AAFTSQFFPSRRSRASPGSAP\GNGQNLTEQHP CPGSCDPQVLSASWM*VEHRSKFRPPP*NSTI PPES/RS*QGGTVQTGQHSSGREAGSWRARGR NAGRR*KGGGKIGTKQGAVRARKECRGEMA SGETDSE
430	1780	A	3473	2802	270	FRMRIFLHCPWNQQMWKIWNLLETSLESCKA HLSIQKLLKER\Q\QLPVFKHRDSIVETLKRHR VVVVAGET\GSGKSTQVPHFLLEDLLLNEWE ASKCNIVCTQPRRISAVSLANRVCDELGCENG PGGRNSLCGYQIRMESRACESTRLLYCTTGV LLRKLQEDGLLSNVS/HMFIVDEV\HER\SVQS DFLLIILKEILQKRSDLHLILMSATVDSEKFST YFTHCPILRISGRSYPVEVFHLEDIIEETGFVLE KDSEYCQKFLEEEEEVTINVTSKAGGIKKYQE YIPVQTGAHADLNPFYQKYSSRTQHAILYMN PHKINLDLILELLAYLDKSPQFRNIEGAVLIFL PGLAHIQQLYDLLSNDRRFYSERYKVIALHSI LSTQDQAAAFTLPPPGVRKIVLATNIAETGITI PDVVFVIDTGRTKENKYHESSQMSSLVETFVS KASALQRQGRAGRVRDGFCFRMYTRERFEG FMDYSVPEILRVPLEELCLHIMKCNLGSPEDF LSKALDPPQLQVISNAMNLLRKIGACELNEPK LTPLGQHLAALPVNVKIGKMLIFGAIFGCLDP VATLAAVMTEKSPFTTPIGRKDEADLAKSAL AMADSDHLTIYNAYLGWKKARQEGGYRSEI TYCRRNFLNRTSLLTLEDVKQELIKLVKAAGF SSSTTSTSWEGNRASQTLSFQEIALLKAVLVA GLYDNVGKIIYTKSVDVTEKLACIVETAQGK AQVHPSSVNRDLQTHGWLLYQEKIRYARVY LRETTLITPFPVLLFGGDIEVQHRERLLSIDGW IYFQAPVKIAVIFCGSSAALGCGLESUPPRGGO
431	1781	A	3474	1	441	FRPAPGHVQP*GGSSAAAGGGLLSHPRPCQQ PCPPAPAPSRPRSLGSLGQRVPAALATAAQEL PATLGGDGGKPALTAGEAALPGLHRSGVPAA AARC*PCT/SRPT*STLSPTQAAWWCRPSRRQ QRGEASTGGASGRRCGSCFQV

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid comence (A-Alexina C-Chatain
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1.00	in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		Į.	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		İ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	١.			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide	•	/=possible nucleotide deletion, \=possible
			1	sequence		nucleotide insertion
432	1782	A	3478	416	23	QLRRLTLPNFKTY/YSS*IIEIAWH**KNMQID
		İ				QWFRRESPEIDLCKYS*LSFDKEAKAIK/WKE
		1				CSLFNKWC/YKNWM/LHVQKKRI*VQTLHPS
	[ĺ	[ĺ	•	QKLK\SKWIKDLNVECRITKLLDQEYPGDLGY
						SRALNSGSR
433	1783	A	3504	1876	552	CLAPCSPQPEKNGMQPLLLLLPPLLYQQLLHS
						SLGAPGESTLLVRTSKLLVGLGLQLLVWLLL
	l	i	ł			QTRSLLALQLHLTSSAPLLAAPTAVCSCSRCS
						APRSRCVARPAARTGLPTPAPASSPAPAASPA
	\		1			PAASPAPAESTA\PQPLILLPKP/PPAPGAPPPRP
	l .		1			GAPPPRPAASPSPAASPAPPAASPVLTASPPLP
	j	1	1 .	j		AASPSPAASPAPPAASPVLTASPPLPAASPSPA
			1			ASPAPPAASPVLTASPPLPAASPALAASPVHT
		l	1			ASPPVHVASPPVHTASPPVHVASPPVHTASPP VHVASPPVHTASPHVHVASPPVHVASPPVHV
					ļ	ASPPVHTASPPVHVASPPVHTASPHVHVASPP
	1		}			VHTASPPVHVASPPVHVASPPVHVAYPPVHV
			İ			ASPPVHVASPPVHVASPPVSCSGDSTSDCFPP
						QPGAVFPHSLAPSLGGWSHLVAALP
434	1784	A	3516	142	590	GGVNRPRSETEQVKTPVLISSWDYRHPPPRPA
		i	ł			SFFVFLV*TGF\TALARMVLISWPCDLPTSASO
			}		1	SAGITGVRHHA\RLLYFEQESHSVTQAGW\VQ
			Ì			WHNLGSLQPLSLEDRLSPGVLGCSALCRSGV
					•	RTKFGINMVTSRERGTTRLPKEG
435	1785	A	3529	1	3161	MSLVRAALEALDELDLFGVKGGPQSVIHVLA
		ļ				DEVQHCQSILNSLLPRASTSKEVDASLLSVVS
		ŀ				FPAFAVEDSQLVELTKQEIITKLQGRYGCCRF
						LRDGYKTPKEDPNRLYY/ENPAELKLFENIEC
			1.		·	EWPLFWTYFILDGVFSGNAEQVQEYKEALEA
		[1			VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT
						VDRVPMGKLPHMWGQSLYILGSLMAEGFLA
			i			PGEIDPLNRRFSTVPKPDVVVQVYPSLPHGCS
	ĺ	1	[SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL
						KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK
		ŀ			}	LYDIRKTIFTFTPQFIDQQQFYLALDNKMIVE
		!				MLRTDLSYLCSRWRMTGQPTTTFPISHSMLDE
	ľ	ŀ				DGTSLNSSILAALRKMQDGYFGGARVOTGKL
	1					SEFLITSCCTHLSFMDPGPEGKLYSEDYDDN
	Ì	1				YDYLESGNWMNDYDSTSHARCGDEVARYL
	1	!				DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT
]	}				TCDLMSLVTKAKELHVQNVHMYLPTKLFQA
		1				SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ
		1				SGEVDFKALVLQLKETSSLQEQADILYMLYT
						MKGPDWNTELYNERSATVRELLTELYGKVG
						EIRHWGLIRYISGILRKKVEALDEACTDLLSH
	}	1				QKHLTVGLPPEPREKTISAPLPYEALTQLIDEA
						SEGDMSISILTQEIMVYLAMYMRTQPGLFAE
						MFRLRIGLIIQVMATELAHSLRCSAEEATEGL
				}		MNLSPSAMKNLLHHILSGKEFGVERSVRPTD
		٠ .				SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK
						QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW
						QRRRRLDGALNRVPVGFYQKVWKVLQKCH
]			GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL
	ļ			1		NRVPQPEYRQLLVEAIL\VLTMLADIENHSIGS
		1]			
						IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD
		×	-		1.0	IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY
426	100		2546		200	IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY VQEFLPHSICAMQ
436	1786	A	3546	73	393	IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD PASGICTLLYDSAPSGRFGTMTYLSKAAATY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN 09/496	location	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		914	correspondi ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uanoo		ł	***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ł		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		j		peptide)	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						KGKGKTIRGI*TFKGRKGGTYQREHDANPLA
437	1000		3554	5157	2939	PXSARSCWMRKG
437	1787	A	3334		2939	AVRAEPGLEELSSGLRAHSPSATTVCEPEAQG SASGCRYAAHPHWGLGGAAAAGGSWEPQPP
				,		RPVCEPAGRGKPHPPAAPRSPLLPGSRRRPHA
						AQPGARARTSPPPASARNMAARPAATLAWSL
					ĺ	LLLSSALLREGCRARFVAERDSEDDGEEPVVF
					ļ	PESPLQSPTVLVAVLARNAAHTLPHFLGCLER
		ĺ	İ		Ì	LDYPKSRMAIWAATDHNVDNTTEIFREWLK
						NVQRLYHYVEWRPMDEPESYPDEIGPKHWP
						TSRFAHVMKLRQAALRTAREKWSDYILFIDV DNFLTNPQTLNLLIAENKTIVAPMLESRGLYS
,					j	NFWCGITPKGFYKRTPDY\VOIREWKRTGCFP
		1	1			VPMVHSTFLIDLRKEASDKLTFYPPHQDYTW
			ļ			TFDDIIVFAFSSRQAGIQMYLCNREHYGYLPIP
		i	İ	ľ	ĺ	LKPHQTLQEDIENLIHVQIEAMIDRPPMEPSQ
		1				YVSVVPKYPDKMGFDEIFMINLKRRKGQGGD
]			1			RWLRTLYEQEIEVKIVEAVDGKALNTSQLKA
				ļ		LNIEMLPGYRDPYSSRPLTRGEIGCFLSHYSV WKEVIDRELEKTLVIEDDVRFEHQFKKKLMK
						LMDNIDQAQLDWELIYIGRKRMQVKEPEKA
					İ	VPNVANLVEADYSYWTLGYVISLEGAQKLV
		1	İ	ľ	Ì	GANPFGKMLPVDEFLPVMYNKHPVAEYKEY
			•			YESRDLKAFSAEPLLIYPTHYTGQPGYLSDTE
		ļ]		1	TSTIWDNETVATDWDRTHAWKSRKQSRIYSN
400	1,500	<u> </u>	200	122		AKNTEALPPPTSLDTVPSRDEL
438	1788	A	3563	130	527	IFFNSSSLFCRVFCLFLRWSFTLVAQARVQ*C
		[ĺ		NLSSLQPLPPGFK*FSCLSPPRS*DYRRPPPRPA NFLYF**RQGFTVLGQAGLELLT/S/GDPPTSA
					1	SOSAGITGVSHRAWPVHAISTHISLVKTRPSLT
						TLG
439	1789	Α	3565	446	1834	LLQPAMRKSPGLSDCLWAWILLLSTLTGRSY
						GQPSLQDELKDNTTVFTRILDRLLDGYDNRL
]	j		RPGLGERVTEVKTDIFVTSFGPVSDHDMEYTI
						DVFFRQSWKDERLKFKGPMTVLRLNNLMAS
			Ì			KIWTPDTFFHNGKKSVAHNMTMPNKLLRITE DGTLLYTMRLTVR\AECPMAFGRDFPM\D\AH
			1	ł		ACPLKFGSYAYTRAEVVYEWTREPARSVVV
				1		AEDGSRLNQYDLLGQTVDSGIVQSSTGEYVV
						MTTHFHLKRKIGYFVIQTYLPCIMTVILSQVSF
		1				WLNRESVPARTVFGVTTVLTMTTLSISARNSL
		[1	1		PKVAYATAMDWFIAVCYAFVFSALIEFATVN
		1]	1		YFTKRGYAWDGKSVVPEKPKKVKDPLIKKN NTYAPTATSYTPNLARGDPGLATIAKSATIEP
		i	1	1	!	KEVKPETKPPEPKKTFNSVSKIDRLSRIAFPLL
				1		FGIFNLVYWATYLNREPQLKAPTPHQ
440	1790	A	3568	1	350	STSSCFPAAAAAIMREIVHLQAGQCGNQIGAK
		[FWEVISDEHGIDPTGTYHGDSDLQLERINVYY
		İ	!	 	}	NEATGEAPVPSPTALRGPRGPCLG*RPPVPAG
						GKYVPRAVLVDMEPGTMDSV
441	1791	A	3569	2	1751	FVAVAGAVSGEPLVHWCTQQLRKTFGLDVS
					1	EEIIQYVLSIESAEEIREYVTDLLQGNEGKKGQ
]]	J		FIEELITKWQKNDQELISDPLQQCFKKDEILDG
			!			QKSGDHLKRGRKKGRNRQEVPAFTEPDTTAE VKTPFDLAKAQENSNSVKKKTKFVNLYTREG
			1	ł		QDRLAVLLPGRHPCDCLGQKHKLINNCLICG
		l	1	l	1	RIVCEQEGSGPCLFCGTLVCTHEEQDILRGDS
			}			N/KSQKLLKKLMSGVENSGKVDISTKDLLPH
1				·		QELRIKSGLEKAIKHKDKLLEFDRTSIRRTQVI

	000 10					
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in			F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi		M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ŀ	ł	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
		ļ		sequence		nucleotide insertion
		l				DDESDYFASDSNQWLSKLERETLQKREEELR
						ELRHASRLSKKVTIDFAGRKILEEENSLAEYH
	i					SRLDETIQAIANGTLNQPLTKLDRSSEEPLGVL
]				VNPNMYQSPPQWVDHTGAASQKKAFRSSGF
						GLEFNSFQHQLRIQDQEFQEGFDGGWCLSVH
			}		į	QPWASLLVRGIKRVEGRSWYTPHRGRLWIAA
			1			TAKKPSPQEVSELQATYRLLRGKDVEFPNDY
		Į	İ			PSGCLLGCVDLIDCLSQKQFKEQFPDISQESDS
	1		ļ		1	PFVFICKNPQEMVVKFPIKGNPKIWKLDSKIH
						QGAKKGLMKQNKAV
442	1792	A	3576	1	2019	MPRSHTGERLCEGKEGSQCAENFSPNLSVTK
						KTAGVKPYECTICGKAFMRLSSLTRHMRSHT
	·		1			AIRANEKPYKCKEC\GRAFSLSQILSK\HERSH
			l			TGEKPYKCKQCGKTFIYHQPFQRHERTHIGEK
		1				PYECKQCGKALSCSSSLRVHERIHTGEKPYEC
	1		1			KQCGKAFSCSSSIRVHERTHTGEKPYACK\EC
		1				GKAFIS\TTSVLTHMITHNGDRPYKCKECGKA
	i	1				FIFPSFLRVHERIHTGEKPYKCKQCGKAFRWS
1	İ			1		TSIQIHERIHTGEKPYKCKECGKSFSARPAFRV
		1	1		•	HVRVHTGEKPYKCKECGKAFSRISYFRIHERT
		1				HTGEKPYECKKCGKTFNYPLDLKIHKRNHTG
1		1				EKPYECKECAKTFISLENFRRHMITHTGDGPY
						KCRDCGKVFIFPSALRTHERTHTGEKPYECKQ
ľ		ľ				CGKAFSCSSYIRIHKRTHTGEK\PYECKECGK
	İ					AFIYPTSFQGHMRMHTGEKPYKCKECGKAFS
					·	LHSSFR\RHTRIHNYEKPLEC*Q\CGKAFSVSTS
						LKKPMRNAQSDRKLY/KCEK*EKVFNSNRCF
				{		QSCENSH*REKSCQCK*YRKRDTR*FMYSQV
		!		1		PHNHVSVSNGPYR/CGSPIRLYNT*NISINRNL
1		1				VAVVTP*CSTLFKCLWCWCKRAALSVV*/IVQ
			İ			DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL
						ANTVKPHLY
443	1793	A	3578	287	114	DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK
		l			L	LSIIKKSVLQNNL*FSAASMRFQKVFF
444	1794	A	3582	3335	1909	HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT
	1	ľ		I	1	MTKVTLENFYSNLIAQHEEREMRQKKLEKV
		1.		1	1	MEEEGLKDEEKRLRRSAHARKETEFLRLKRT
	1			i	1	RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH
				-	1	VYAMKILRKADMLEKEQVGHIRAERDILVEA
		1		1	1	DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM
		1		1	1	MTLLMKKDTLTEEETQFYIAETVLAIDSIHQL
		1		1		GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK
				1	1	KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE
				1	1	TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN
				1	1	KLCDWWSLGVIMYEMLIGYPPFCSETPQETY
		1		[KKVMNWKETLTFPPEVPISEKAKDLILRFCCE
1	1	1		1		WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA
	1	1		1		AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE
1				I	I	TDYKNKDWVFINYTYKRFEGLTARGAIPSYM
}			ļ		1	KAAK
445	1795	A	3584	1	6169	RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD
ļ	1	_				GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK
1					1	NHRLYFLSAASRPLCSGGHASNKEKEMVTSL
1		1		l		FCKLGVLVRHRISLFGNDATSIVNCLHILGQT
				1	1	LDARTVMKTGLESVKSALRAFLDNAAEDLE
	1					KTMENLKQGQFTHTRNQPKGVTQIINYTTVA
l	ľ	1	1		12	LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI
1			1	I	I	
		i	ł	1		LTSLYALGTSKSIYVERORSALGECLAAFAGA
			;			LTSLYALGTSKSIYVERQRSALGECLAAFAGA FPVAFLETHLDKHNIYSIYNTKSSRERAAISIP
						LTSLYALGTSKSIYVERQRSALGECLAAFAGA FPVAFLETHLDKHNIYSIYNTKSSRERAALSLP TNVEDVCPNIPSLEKLMEEIVELAESGIRYTQ

CEO ID	SEO ID	3.6-+	CEC	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ			Amino acid sequence (A-Atalinie C-Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutarnic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i i				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
					sequence	/=possible nucleotide deletion, \=possible
į.	Ì			peptide		
L				sequence		nucleotide insertion
1	1					MPHVMEVILPMLCSYMSRWWEHGPENNPER
İ						AEMCCTALNSEHMNTLLGNILKIIYNNLGIDE
ł	i i					GAWMKRLAVFSQPIINKVKPQLLKTHFLPLM
1	J .					EKLKKKAATVVSEEDHLKAEARGDMSEAEL
1						LILDEFTTLARDLYAFYPLLIRFGDYNRAKWL
						KEPNPEAEELFRMVAEVFIYWSKSHNFKREE
1						
1						QNFVVQNEINNMSFLITDTKSKMSKAAVSDQ
1	ļ					ERKKMKRKGDRYSMQTSLIVAALKRLLPIGL
1	1					NICAPGDQELIALAKNRFSLKDTEDEVRDIIRS
1						NIHLQGKLEDPAIRWQMALYKDLPNRTDDTS
i		l				DPEKTVERVLDIANVLFHLEQKSKRVGRRHY
ì		1				CLVEHPORSKKAVWHKLLSKORKRAVVACF
1	}	ł			1	RMAPLYNLPRHRAVNLFLOGYEKSWIETEEH
				· ·		YFEDKLIEDLAKPGAEPPEEDEGTKRVDPLHQ
1	1	1				,
}	1	1			}	LILLFSRTALTEKCKLEEDFLYMAYADIMAKS
J	ľ			j		CHDEEDDDGEEEVKSFEEKEMEKQKLLYQQ
	i				ł	ARLHDRGAAEMVLQTISASKGETGPMVAAT
ľ	1					LKLGIAILNGGNSTVQQKMLDYLKEKKDVGF
i	İ					FQSLAGLMQSCSVLDLNAFERQNKAEGLGM
1	1	į				VTEEGSGEKVLODDEFTCDLFRFLOLLCEGH
ĺ	ĺ	1			[NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVQ
1	!					ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA
ì	i				Ι.	
ĺ	ļ					KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV
1	1	}			ł	VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ
j	ł	į			1	KDMVVMLLSMLEGNVVNGTIGKQMVDMLV
-						ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD
l .						PDGKGVIFKRDFHKAMESHKHYTQSETEFLL
1	1	ļ			}	SCAETDENETLDYEEFVKRFHEPAKDIGFNVA
1	1		İ			VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP
f	[•			[FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ
1		ł				VKESKROFIFDVVNEGGEKEKMELFVNFCED
	}	l	1			
1				ł	ļ	TIFEMQLAAQISESDLNERSANKEESEKERPEE
1	1	[ĺ	QGPRMAFFSILTVRSALFALRYNILTLMRMLS
1	ł	ł		1		LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI
1						FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA
1		l ·				KKIKVAELLANMPDPTQDEVRGDGEEGERKP
1	ì	1			1	LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR
1.	1	1)	ļ		EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE
1		1				KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE
1	1	[KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA
1	1	l]	FWKKIIAYQQKLLNYFARNFYNMRMLALFV
1	}]	Ì	1		
		}			1	AFAINFILLFYKVSTSSVVEGKELPTRSSSENA
í		[-	[[KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL
		İ				PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK
	!					LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP
1	1	1	[1	NNYWDKFVKRKVMDKYGEFYGRDRISELLG
ļ	1	1	}	1	ł	MDKAALDFSDAREKKKPKKDSSLSAVLNSID
1	1	1		l	1	VKYQMWKLGVVFTDNSFLYLAWYMTMSVL
İ		1		1	1	GHYNNFFFAAHLLDIAMGFKTLRTILSSVTH
1	1	l	1	l	1	NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF
	1	1	}	ł		
	1]		}]	YNKSEDGDTPDMKCDDMLTCYMFHMYVGV
	i	1			1	RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI
		1			1	VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC
}	1	J	ļ	}	1	FICGIGNDYFDTVPHGFETHTLQEHNLANYLF
						FLMYLINKDETEHTGQESYVWKMYQERCWE
1	Í	ĺ		1	1	FFPAGDCFRKQYEDQLN
446	1796	A	3592	1	355	AGLELLNSDDPPALASQSAGITGVTRTPSLFF*
1 770	1	١.,		ļ [~]	333	DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL
i	1	۱.	{	[1	SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML
1	1	1	ľ	(ł	
L	l	L	L	L	L	PRLVSNSWTQAILLPRPPKMLGLQV

070 70	CEO ID	1 X Cat	CEO	Dradiated	Doodisted and	Amino soid
SEQ ID	SEQ ID	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of nucl-	NO: of peptide	1100	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	l	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
uence		ļ	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ		Ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	•		peptide	scquence	/=possible nucleotide deletion, \=possible
		l		sequence		nucleotide insertion
447	1797	A	3598	1202	1070	LFVGGGPICPEGASGFAPGPAPAPRVGVDAEV
447	1191	^	3376	1202	1070	GR*V*GAAASQGA/GSLRPRPTGPGHPGAWL
						QVWGAAAVCAGPAM*/AVRAKRGPRAG*EP
		j		ļ	,	NSPWRSGVLAA\RAVGAGPWP*P*PGCS*ARG
						PSSRSAPGLASGPAAPLLQGVHSSAGPLLCYI
	ĺ			1		
140	1798	<u> </u>	3604	3115	557	NGTLALGLKP**AWGWGEWRPKG
448	1/98	Α	3004	3112	33/	FRRKGGGGPKDFGAGLKYNSRHEKVNGLEE
		}	į			GVEFLPVNNVKKVEKHGPGRWVVLAAVLIG
	j	}	ļ]		LLLVLLGIGFLVWHLQYRDVRVQKVFNGYM
	1		Ì			RITNENFVDAYENSNSTEFVSLASKVKDALKL
		ł				LYSGVPFLGPYHKESAVTAFSEGSVIAYYWSE
	1	1				FSIPQHLVEEAERVMAEERVVMLPPRARSLKS
	1	1	1	l		FVVTSVVAFPTDSKTVQRTQDNSCSFGLHAR
				1		GVELMRFTTPGFPDSPYPAHARCQWALRGD ADSVLSLTFRSFDLASCDERGRHLV\TVYNT\L
	1	ļ	}			
	}		}]	SPMEPHA\LVQLCGTYPPSYNLTFHS\S\QNVL
			1			LITLITNTERRHPG/FEATFFQLPRMSSCGGRL RKAQGTFNSPYYPGHYPPNIDCTWNIEVPNN
		1	l	١.		QHVKVRFKFFYLLEPGVPAGTCPKDYVEING
	(1		Í	1	EKYCGERSOFVVTSNSNKITVRFHSDOSYTDT
	İ	İ				
			}		ļ	GFLAEYLSYDSSDPCPGQFTCRTGRCIRKELR CDGWADCTDHSDELNCSCDAGHQFTCKNKF
]	})	
			•			CKPLFWVCDSLNDCGDNSDEQGCSCP\AQTF
		l			Ì	RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK
	Ì	ĺ	l	1	1	EDCSDGSDEKDCDCGLRSFTRQARVVGGTD
	ļ	}]		}	ADEGEWPWQVSLHALGQGHICGASLISPNWL
)	ļ	VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS
			1		i	QRSAPGVQERRLKRIISHPFFNDFTFDYDIALL
		٠ .				ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV
		1	1	1	ĺ	TGWGHTQYGGTGALILQKGEIRVINQTTCEN
						LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL
		ĺ	ł	1		SSVEADGRIFQAGVVSWGDGCAQRNKPGVY
449	1799	<u> </u>	3618	2	712	TRLPLFRDWIKENTGV
449	1/99	A	3018	2	613	FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ
		1			1	EMTRRPSLMAGRQHGWSAQQSATVANPVPG
		[Í		ĺ	ANPOLLPHFLGEPEDVYIVKNKPVLLVCKAV
		1		1		PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP
		1	1	1	1	TMEVRINVSRQQVEKVFGLEEYWCQCVAWS
		1	l	1	1	SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL
450	1000	 	2600	 	2026	EQGIVLPCRPPEGIPPAE
450	1800	A	3620	1	2676	MEPSLGQGMDLTCPFGVSPACGAQASWSIFG
		ł	!	1		ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE
				1		PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA
		1	1	}		TSPEGETDKNLANRVHSPHKRLSHRHLKVST
		1	1	1	1	ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN
		1	!			LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA
	}	j.	}			VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS
	1	[1	SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP
	1	i	1			RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA
	1	1				PVTNSSGKMALNSPQPGPVESELGKQLLKTG
	[1	1	1		WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP
		1	1	1		MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE
I	1	1			1	PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK
\	(í	1	Į.		APGLKDFQIQVQPVRMQKLTKLREEHILMRN
		I .			t .	QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE
				1	ŀ	
				}		ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP
						ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE
	-					ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid residue of	corresponding to last amino acid residue of peptide sequence	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
						VRQEKRMSKATEVMMQYVENLKRTYEKDH AELMEFKKLANQNSSRSCGPSEDGVLRTARS MSLTLGKNMPRRRVSVAVVPKFNALNLPGQ TPSSSSIPSLPALSESPNGKGSLPVTSALPALLE NGKTNGDPDCEASAPALTLSCLEELSQETKA RMEEEAYSKGFQEGLKKTKELQDLKEEEEEQ KSESPEEPEEVEETEEEEKDPRSSKLEELVHFL QVMYPKLCQHWQVIWMMAAVMLVLTVVL GLYNSYNSCAEQADGPLGRSTCSAAQKDSW
				·		WSSGLQHEQPTEQ
451	1801	A	3623	504	198	QLIQHQTVHTGRKLYECKECGKAFNQGSTLI RHQRIHTGEKPYECKVCGKAFRVSSQLKQHQ RIHTGERPYQCKELKGRGAEMLAVLAVKEQ
452	1802	A	3628	2	195	NRTPVNYGK MTCLHSAKAFHY*SSCSFSCEEGFALIGPEVV QCTALGVWTAPAPVCIAVQCQHLEALNEGT MG*DYPFTAFAYGSSCKYECHTVYRVRGLD MLHSRGCYLWNGHFTT*EAISCEPLERPCH*S V*CSFSCEEGFALIGPEVVQCTALGVWTAPAP VCIAVOCOHLEALNEGTMG
453	1803	A	3637	662	142	IQAKGLGIWHVPNKSPMQHWR\KGSLLRYRT DTGFLQTLGHNLLGIYQKYPVKYGEGKCWT DNGPVIPVVYDFGDAQKTASYYSPYGQREFT AGFVQFRVFNNERAANALCAGMRVTGCNTE HHCIGGGGYFPEASPQQCGDFSGFDWSGYGT \HVGYSSSREITE\AAVLLFYR
454	1804	A	3641	I	362	TQVHPAMLGLDELGRSGCGHCTQADLRFGD AAGRDPGQDNDRNTAEPAFPPPPRVMAAAA ALRAPAQSSVTFEDVAVNFSLEEWSLLNEAQ GCLYHDVMLETLTLISSLGKVLILNCDLS
455	1805	A	3646	2	414	AAAGRGASGALTGEGGGEQGRRVGLGSRAH SLLLGPTFNSCQVSSQPPRVAGLGLPLKHEPS RPQPPSPRGPRTVRAGVPGAHPQDTPCPEFVR PRKVPLVGEAPGLPPEERSRGWRRDTPGLQE SRVRAPSYDDIT
456	1806	A	3656	396	8	QIVSFNSYLTLYTKNNLKSMKDLNVNTEMIK LLELKNIHNLG*AKFFLN*IQKALIKRKILIHW P/LIKIK/SFCSLSDTIKKMKRQTIVWEQTFIIHI SVKELVSRIYEAFLQFNKTVNRPVFDIKKEQK F
457	1807	A	3660	14	1961	SEAKLGGPTGMDLWQLLLTLALAGSSDAFSG SEATAAILSRAPWSLQSVNPGLKTNSSKEPKF TKCRSPERETFSCHWTDEVHHGTKNLGPIQLF YTRRNTQEWTQEWKECPDYVSAGENSCYFN SSFTSIWIPYCIKLTSNGGTVDEKCFSVDEIVQ PDPPIALNWTLLNVSLTGIHADIQVRWEAPRN ADIQKGWMVLEYELQYKEVNETKWKMMDP ILTTSVPVYSLKVDKEYEVRVRSKQRNSGNY GEFSEVLYVTLPQMSQFTCEEDFYFPWLLIIIF GIFGLTVMLFVFLFSKQQRIKMLILPPVPVPKI KGIDPDLLKEGKLEEVNTILAIHDSYKPEFHS DDSWVEFIELDIDEPDEKTEESDTDRLLSSDH EKLHINLGVKDGDSGRTSCCEPDILETDFNAH DIHEGTSEVAQPQRLKGEADLLCLDQKNQNN SPYHDACPATQQPSVIQAEKNKPQPLPTEGAE STHQAAHIQLSNPSSLSNIDFYAQVSDITPAGS VVLSPGQKNKAGMSQCDMHPEMVSLCQENF LMDNAYFCEADAKKCIPVAPHIKVESHIQP\S

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ì	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide	•	/=possible nucleotide deletion, \=possible
•		1		sequence	Į	nucleotide insertion
		t	 	 	 	PVPDYTSIHIVQSPQGLILNATALPLPDKEFLS
		İ			1	SCGYVSTDQLNKIMP
458	1808	A	3663	154	462	TRAPASGRSGAGLALSANAPDSGGHPGATEG
						PAGSLAHASGSARGTWRVRGRGSHGWERTV
		i				GAGGCANPVPALHSCASAPRGTGRVSALGPK
		[1	1	ĺ	TGSSPLSSPKG
459	1809	A	3664	902	135	LGKYNTSMALFDFVLHNSTGEIRYITEDDVIQ
,						SQNALGKYNTSMALFESNSFEKTILESPYYVD
		1	i		<u> </u>	LNQTLFVQVSLHTSDPNLVVFLDTCRASPTSD
		ĺ		Ī		FASPTYDLIKSGCSRDETCK\VYPLFGHYGRF
		i		1		QFNAFKFLRSMSSVYLQCKVLICDSSDHQSRC
		l		1	1	WQGCVSRSKRDISSYKWKTDSIIGPIRLKRDR
		1	1			SAINGNSGFQHETHAEETPNQPFNSVHLFSFM
		1				VLALNVVTVATITVRHFVNQRADYQ\YQKLQ
		1				NY
460	1810	A	3670	850	557	LGILMSPQVEAGEI*ALLTPPPGCMQFSPLTL/P
400	1010] ^	3070	850) 337	K*WVSPGLTP/PPPEVPSVFLVEPGLPHAGQA
					ļ	GLDLL\TSGDPPASTSQSARTTDVSHRAQPLAI
						S
461	1811	A	3671	2472	2099	IGVLAFETGSCSVTRLYCIGIIMPHCSLDLAGS\
401	1011	\ ^	30/1	2412	2099	
		1			1	TSAFRIAGTTSVHHHPQLTFFFFWIETGSHCV
		i				VQTGL*LLALSNPPALASQIAGISGMSHRAWP
160		 	1000			GLVLYSLEFSLLCASQSLIMLFTCYNE
462	1812	Α	3672	394	110	VKPVNGESKRD*GADTQTCEGEADEQLQT\N
		1				CYYD/STKSFFYISCG*K\RKPTWAENRRLNA
						KMFGIPLHSNSDPWGYEEREVIGFHRSRVSRG
162	1012	 	2.00		 	HGS
463	1813	A	3673 .	348	1	QRNPFSAGHPQRPPTSGSQSELLAQPRLRPGR
		1	1		Ì	KSSFSRDQDVW*SQAVPKRQ*QRNPFSAGHP
	ļ	ĺ]	QRPPTSGSQSELLAQPRLRPGRKSSFSRDQDV
					·	WPGQKPRPSQQQHQMCASPTLGQRSPFALEP
				1		VPAYHGGRDPFASARPSPVGIPKPRAAPAGG
						GWRRIRPKSSTK
464	1814	Α	3676	2253	320	PVIQRCSQPYGFSLLISFFLKCVSETSQQPPSR
		1				KVFQLLPSFPTLTRSKSHESQLGNRIDDVSSM
	}					RFDLSHGSPQMVRRDIGLSVTHRFSTKSWLS
	ļ	}	į		}	QVCHVCQKSMIFGVKCKHCRLKCHNKCTKE
					}	APACRISFLPLTRLRRTESVPSDINNPVDRAAE
		!	1		į	PHFGTLPKALTKKEHPPAMNHLDSSSNPSSTT
	1	١.		ſ		FSTPSSPAPFPTSSNPSSATTPP\NPSP\GQR\DSR
						FNFPSC/AYFIHHR\Q\QFIFPDISAFAHAAPLPE
	}	İ	l	Į.		AADGTRLDDQPKADVLEAHEAEAEEPEAGK
	Į.		1	1		SEAEDDEDEVDDLPSSRRPWRGPISRKASQTS
	1	1	[1	ĺ	VYLQEWDIPFEQVELGEPIGQGRWGRVHRGR
	ł	{	1		1	WHGEVAIRLLEMDGHNQDHLKLFKKEVMN
	!	1				YROTRHENVVLFMGACMNPPHLAIITSFCKG
			1	1	1	RTLHSFVRDPKTSLDINKTRQIAQEIIKGMGY
	1	1	[i	LHAKGIVHKDLKSRNVFYDNG\KVVITDFGLF
	1	1	1	1	1	GISGVVP\EGRRENQLKLSHDWLCYLAPEIVR
			1		1	EMTPGKDEDQLPFSKAADVYAFGTVWYELQ
	!	ļ	1	į.		ARDWPLKNQAAEASIWQIGSGEGMKRVLTS
	1	ı	1	[VSLGKEVSENLSACWAFDLQERPS\FSLLMD
	l	1	1	1		MI EKI DKI MDDI CHDOLIDAKO A DDIGORIA DDI
	1	}	1	1	1	MLEKLPKLNRRLSHPGHF*KSADINSSKVVPR
100	1015	 	2/20	 	L.002	FERFGLGVLESSNPKM
465	1815	A	3679	8	803	IPSPAWWNSTWADTFSLLLALAVALYLGYY
	1	1	1	1	1	WACVLQTHRAFCASNTEDLETVVNHIKHRYP
1)	i	1	I	1	QAPLLAVGISFGGILVLNHLAQARQAAGLVA
	l	1	ì	1	ł	The state of the s
						ALTLSACWDSFETTRSLETPLNSLLFNQPLTA GLCQLVERLSY/E*DLQARTIRQFDERYTSVA

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid residue of	corresponding to last amino acid residue of peptide sequence	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence	sequence	/=possible nucleotide deletion, \=possible nucleotide insertion
						FGYQDCVTYYKAASPRTKIDAIRIPVLYLSAA DDPFSTVCALPKQAAQHSPYVALLITARGGHI GFLEGLLPWQHWYMSRLLHQYAKAIFQDPE
466	1816	A	3684	3	307	GLPDLRALLPSEDRNS SSQYIVQSKTKIFL*AAREKQ/RHTCRRFSIRLS ANISSQTGEARGQWPSVFKVLKEKKLSTKKS FGQK*GR\RKTFPDKQK/LREFDTTRPTIQEML TGVLQG
467	1817	A	3687	2465	837	ELPTPLIAAHQLYNYVADHASSYHMKPLRMA RPGGPEHNEYALVSAWHSSGSYLDSEGLRHQ DDFDVSLLVCHCAAPFEEQGEAERHVLRLQF FVVLTSQRELFPRLTADMRRFRKPPRLPPEPE APGSSAGSPGEASGLILAPGPAPLFPPLAAEVG MARARLAQLVRLAGGHCRRDTLWKRLFLLE PPGPDRLRLGGRLALAELEELLEAVHAKSIGD IDPQLDCFLSMTVSWYQSLIKVLLSRFPQSCR
		-				HFQSPDLGTQYLVVLNQKFTDCFVLVFLDSH LGKTSLTVVFREPFPVQPQDSESPPAQLVSTY HHLESVINTACFTLWTRLL*GSGLDH*MSLFL ESWAYQIACQRQD*PALLGPRASQTLSDTKG FVTMS*GSAAPAWQQEPPSPNTHSH*PIQDSR ESGQPRGPLGPFWGTPFGPPGRVSGYHTGWQ TPPRAPLPESCPL\PLTTVSHLCPLSLRVFTSHL DITAGHSHRDDTWVPIPALPLKHLRPPSSPFA LGPWVSHPLMRWVQKLSHLHSNPGTGFSMG GKQQRN
468	1818	А	3691	960	499	OTCRKDKRAIYPHFONE*MNEIKAI*SGTGGI OCFHSONDSAFFFFLFLLETEFCSAA/TVOWH DFLSMQPPPPGFKQFTCLSLLSSWNYRR\PPPF PGNF*FLVKTGFPHVGQTGFELLTSSDLAPLA SONGGITGMSPCAWPFFFFFFGLC
469	1819	A	3714	4747	495	MAYSWOTDPNPNESHEKQYEHQEFLFVNQP HSSSQVSLGFDQIVDEISGKIPHYESEIDENTFF VPTAPKWDSTGHSLNEAHQISLNEFTSKSREL SWHQVSKAPAIGFSPSVLPKPQNTNKECSWG SPIGKHHGADDSRFSILAPSFTSLDKINLEKEL
	!					ENENHNYHIGFESSIPPTNSSFSSDFMPKEENK RSGHVNIVEPSLMLLKGSLQPGMWESTWQK NIESIGCSIQLVEVPQSSNTSLASFCNKVKKIR ERYHAADVNFNSGKIWSTTTAFPYQLFSKTK FNIHIFIDNSTQPLHFMPCANYLVKDLIAEILH FCTNDQLLPKDHILSVWGSEEFLQNDHCLGS
						HKMFQKDKSVIQLHLQKSREAPGKLSRKHEE DHSQFYLNQLLEFMHIWKVSRQCLLTLIRKY DFHLKYLLKTQENVYNIIEEVKKICSVLGCVE TKQITDAVNELSLILQRKGENFYQSSETSAKG LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR
						INFPLEIKSLPRESMLTVKLFGIACATNNANLL AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW
				·	·	DERTVSEMHTILRRWTFSQPLEALGLLTSSFP DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV KFEWNLESPLVQLLLHRSLQSIQVAHRLYWL LKNAENEAYFKSWYQKLLAALQFCAGKALN DEFSKEQKLIKILGDIGERVKSASDHQRQEVL KKEIGRLEEFFQDVNTCHLPLNPALCIKGIDH
			L	<u> </u>		DACSYFTSNALPLKITFINANLMGKNISIIFKA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, =possible nucleotide deletion, \=possible nucleotide insertion
					·	GDDLRQDMLVLQLIQVMDNIWLQEGLDMQ MIIYRCLSTGKDQRLVQMVPDAVTLAKIHRH SGLIGPLKENTIKKWFSQHNHLKADYEKALR NFFYSCAGWCVVTFILGVCDRHNDNIMLTKS GHMFHIDFGKFLGHAQTFGGIKRDRAPFIFTS EMEYFITEGGKNPQHFQDFVELCCRAYNIIR KHSQLLLNLLÆMMLYAGLPELSGIQDLKY VYNNLRPQDTDLEATSHFTKKIKESLECFPVK LNNLIHTLAQMSAISPAKSTSQTFPQESCLLST TRSIERATILGFSKKSSNLYLIQVTHSNNETSL TEKSFEQFSKLHSQLQKQFASLTLPEFPHWW HILPFTNSDHRRFRDLNHYMEQILNVSHEVTN SDCVLSFFLSEAGQQTVEESSPVYLGEKFPDK KPKVQLVISYEDVKLTILVKHMKNIHLPDGSA PSAHVEFYLLPYPSEVRRRKTKSVPKCTDPTY NEIVVYDEVTELQGHVLMLIVKSKTVFVGAI NIRLCSVPLDKEKWYPLGNSII*PLLLFSSFGM KSLEKDEFVGGMILSNPIW
470	1820	Α	3718	430	75	SHGSISILNLHQGCVFLPSLPAQGLRCYRCLA VLEGASCSVVSCPFLDGVCVSQKVSV/CWQ*/ CPWGARAEGRLSAVVDSQISCCKGDLCNAV VLAAGSPWALCVQLLLSLGSVFLWALL
471	1821	A	3723	891	494	LRQSL/NSVPQAGVQWRDSSLQAPPPRFTPLS CLSLPSSWDYRRLPPCLANFLYF**RRGFTML ARMVLIS*PRDPPASASQ\STETI GGSHRAQHP TDSRDHSERSVKKSHEVISELRMKVIKCKVAF SKNPI
472	1822	A	3734	443	251	GFIET*NFCVSKDTSKKLS/RLPTKWKNVFAN
473	1823	A	3746	3	500	*ISDKGLVSRICQELLRHLDAEQVSSTAGLSL THASGGARSGAGWAGRGVRAGTEAGRGGIF LTLSILRTRDLPSGAMSEGVDLIDIYADEEFNQ DPEFNNTDQIDLYDDVLTATSQPSDDRSSSTE PPPPVRQEPSPKPNNKTPAILYTYSGLRNRRA AVYVGSFSWWTTDQQLIQVIRSIGVYDVGEV KFAENRAK
474	1824		3753		5262	RPLFAREGGIYAVLVCMQEYKTSVLVQQAG LAALKMLAVASSSEIPTFVTGRDSIHSLFDAQ MTREIFASIDSATRPGSESLLLTVPAAVILMLN TEGCSSAARNGLLLLNLLLCNHHTLGDQIITQ ELRDTLFRHSGIAPRTEPMPTTRTILMMLLNR YSEPPGSPERAALETPIIQGQDGSPELLIRSLV GGPSAELLLDLERVLCREGSPGGAVRPLLKRL QQETQPFLLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWLL VHVEPPPGPSPEPSTRPFSKNSKGRDRSPAPSP VLPSSSLRNITQCWLSVVQEQVSRFLAAAWR APDFVPRYCKLYEHLQRAGSELFGPRAAFML

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	nou	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	•		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ĺ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
		<u>-</u>	 	Sequence		ALRSGFSGALLQQSFLTAAHMSEQFARYIDQ
]					QIQGGLIGGAPGVEMLGQLQRHLEPIMVLSG
					'	LELATTFEHFYQHYMADRLLSFGSSWLEGAV
		1				LEQIGLCFPNRLPQLMLQSLSTSEELQRQFHLF
						QLQRLDKLFLEQEDEEEKRL*EEEEEEEEA
		1				EKELFIEDPSPAISILVLSPRCWPVSPLCYLYHP
	ļ	Į	ļ			RKCLPTEFCDALDRFSSFYSQSQNHPVLDMG
		1	ŀ			PHRRLQWTWLGRAELQFGKQILHVSTVQMW LLLKFNOTEEVSVETLLKDSDLSPELLLQALV
						PLTSGNGPLTLHEGQDFPHGGVLRLHEPGPQ
			1			RSGEALWLIPPQAYLNVEKDEGRTLEQKRNL
]			LSCLLVRILKAHGEKGLHIDQLVCLVLEAWQ
1		}				KGPNPPGTLGHTVAGGVACTSTDVLSCILHLL
	1					GQGYVKRRDDRPQILMYAAPEPMGPCRGQA
						DVPFCGSQSETSKPSPEAVATLASLQLPAGRT
		1				MSPQEVEGLMKQTVRQVQETLNLEPDVAQH
	1		!			LLAHSHWGAEQLLQSYSEDPEPLLLAAGLCV
ľ		ł				HQAQAVPVRPDHCPVCVSPLGCDDDLPSLCC MHYCCKSCWNEYLTTRIEONLVLNCTCPIAD
						CPAQPTGAFIRAIVSSPEVISKYEKALLRGYVE
						SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK
						CGWASCFNCSFPEAHYPASCGHMSQWVDDG
						GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE
					-	KNEGCLHMTCAKCNHGFCWRCLKSWKPNH
						KDYYNCSAMVSKAARQEKRFQDYNERCTFH
		ŀ				HQAREFAVNLRNRVSAIHEVPPPRSFTFLNDA
1		}				CQGLEQARKVLAYACVYSFYSQDAEYMDVV
ļ						RLLRADCLSTGMELLRRIQERLLAILQHSAQD
						FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP
	1					EAEEEEDDEDDVPEWQQDEFDEELDNDSFS
						YDESENLDQETFFFGDEEEDEDEAYD
475	1825	Α	3754	1093	96	GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ
						ATGRRRRTRTQQRTAALLTDGTTKTGAAW
						SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN
1				,		PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP
		1				DGTR\RPASITGVAQSPATRATPSLPCLHVPAP SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA
	1		1		,	RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE
	1					PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS
	1			•		HVYIIRATINSISHPLCRAQSSPWEAAGVWRR
	1	l				PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN
						TLWEEGRQRPPETLQPAR
476	1826	A	3758	901	521	FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK
	1					RFSYLSLPSSW\DYRHVPPRQANFCIF/M*RRG
			1			FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI
477	1827	A	3761	843	575	GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG
7''	102/	^	3,01	54.5	3,3	ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI
		1				RPRRPLKVLGLQACTRARLPSPLKEL
478	1828	A	3763	267	1240	HLLSFHLWSASLDCLEQLSQERHVKGMLLGP
		 -				PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS
	[[1			PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P
						GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY
		ŀ	1	* •		WTGHSFASQAWLRQVPEVSKHLQCPSAESLL
			<u> </u>			TMEYHQPEDPAPGKAGTAEAVIPENHEVLAG
ļ]	1	J			PDEHPQDTDARDADGEAREREP/RRPSFAA*P VWGQP\ESPLPEASSAPPGPTLGTLPEVETIRA
	1	İ				CSMPQELP*SPRTRQPEPDFYCVKWIPWKGE
		1				OTPIITQSTNGPLPSPCHHEHPLSSVEGEAPPA
	٠					

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
dence			1 714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of		
1 '	}	ł			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
<u></u>				sequence		nucleotide insertion
						EGSDHIG
479	1829	A	3766	2	2152	YSPIRLLEVCVPLPKIFIKRQAPLKVSLLQDLK
						DFFQKVSQVYVAIDERLASLKTDTFSKTREEK
1						MEDIFAQKEMEEGEFKNWIEKMQARLMSSS
						VDTPQQLQSVFESLIAKKQSLCEVLQAWNNR
						LQDLFQQEKGRKRPSVPPSPGRLRQGEESKIS
[AMDASPRNISPGLQNGEKEDRFLTTLSSQSST
		İ				SSTHLQLPTPPEVMSEQSVGGPPELDTASSSE
1 1		i				DVFDGHLLGSTDSQVKEKSTMKAIFANLLPG
						NSYNPIPFPFDPDKHYLMYEHERVPIAVCEKE
1		1				PSSIIAFALSCKEYRNALEELSKATQWNSAEE
		Ì				GLPTNSTSDSRPKSSSPIRLPEMSGGQTNRTTE
[]				TEPQPTKKASGMLSFFRGTAGKSPDLSSQKRE
]						TLRGADSAYYQVGQTGKEGTENQGVEPQDE
1						VDGGDTQKKQLINPHVELQFSDANAKFYCRL
}						YYAGEFHKMREVILDSSEEDFIRSLSHSSPWQ
						ARGGKSGAAFYATEDDRFILKQMPRLEVQSF
				-		LDFAPHYFNYITNAVQQKRPTALAKILGVYRI
1						GYKNSQNNTEKKLDLLVMENLFYGRKMAQ
[VFDLKGSLRNRNVKTDTGKESCDVVLLDENL
1						LKMVRDNPLYIRSHSKAVLRTSIHSDSHFLSS
						HLIIDYSLLVGRDDTSNELVVGIIDYIRTFTWD
						KKLEMVVKSTGILGGQG*MPTVVSPELYRTR
						FCEAMDNYFLMVPDHCTGLGLNC
480	1830	A	3777	251	3	
700	1050	1	3111	231	,	QGCGSAGTLIHY**ECKMVQLLWKTV*QFLI
				,		KLNI/KDPAITLDVYPNEVKNYVRTKTYTQMF
481	1831		3779	222		I/ANFIMAKSWKQPTHPSVRT
461	1831	A _.	3/19	333	3	EAAIRQPEPNILDVNQIFKDLAMIIHDQGDLID
1						SIEANAESSEVLVERAPGQLQRPA\YYQKKSR
						KKMCLVVLVQTAIILICERIM*VVYTTKWSPPI
						VLPVSCFQGQKFN
482	1832	Α	3780	2	371	TGGRQGKNDHTSITEKPSRDFNRHLITQNI*M
1						PNQDMKSSSNSLIIRKVQIKPTILYHHIFTRKA
						KMKTTDKTKYR*GFKAITTLIHCSQDCKLQ*S
1 1						/L*ENHFMIFPKAEQHITYDTTIPFLR
483	1833	A	3787	43	448	LMKDLSPYVMETHYILNRLNER/RSMWRHIIG
					,	KLPNTKDQEKILKAIRGRREVIQGS/RQQYRR
!						PAAFSAAEKARRLWCS/VFNIERRNL/CEYPTK
1						LSFNIKGEMTFSDKTEFTTNRPSLKMLLKDRI
						QEEGKMF*KEKCFKRKE
484	1834	A	3798	1	727	
"	1037	n.	3170	•	121	FFFFETESRSVAQAGVQWCNLGSLQALPPGF\
1						SHSPASASRVAGTTGTRH*ARLIFYIFSRDGVS
						PC*PGWS*SPDLVIRPP\RLPKCWDYRREPPRP
					,	A*FFVFLVE\QGFTMLARMVSIS*PQ/CDLPAS
j						VSQNAGITGVSHCAWPCLHFCFFGFFFEMESC
						SVAQAEVQWHDLRSLQAPPPGFTPFSCLSLPG
						SWDYRRPPPRPANF\CIFSRDGVSPC*PGWSRS
	<u> </u>	<u> </u>				PDLVIRPPRPPKVLGLQA
485	1835	Α	3802	1	239	FFFFEMECLTVSQAGVQWYNLHSLQPLPPGF
]						KQFSC\LSLPSSWD*RVPTSRPAKF/CVIF*DGV
						SHCOPGWSAVVOPPLH
486	1836	A	3811	378	98	RYD*SSQSENIP\QKEFLLKYP*CTATLGMRN
		**		3.0	70	MSIMKKKSIFSAEFYKVSLPSLLL\HLLAIEWG
						INDIVIDUAL DATE IN A STROTT FULL THE TARK OF
487	1077	<u> </u>	2014	771	330	FHIEIQLTIHQHFLNYELESDFVHIVEYM
40/	1837	A	3814	771	320	FDPDWTRAAGIRHEKKPKALAYRRENSPGDL
						PPPPLPPPEEEASWAL/GAEGSRQHVLPGAGA
						QWGEESGPGRAPGSPAGAPPR*RGLAP\NSRP
						SFLSRGQGTSTCSTAGSNSSRGSSSSRGSRGPG
<u></u>	L					RSRSRSQSRSQSQRPGQKRREEPR

	1 000 10	1 77-7	Lero	D., 31-4-3	F 50 12 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alainne C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	j	ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
İ	Ì		1	peptide		/=possible nucleotide deletion, \=possible
	1	ļ		sequence		nucleotide insertion
488	1838	A	3818	1	781	FRACLLELIPYAPTLSWTACPPAMAGPRGLLP
400	1030	1) 5610	1 1	1 '81	LCLLAFCLAGFSFVRGOVLFKGCDVKTTFVT
İ	1					HVPCTSCAAIKKQTCPSGWLRELPDQITQDCR
				ŀ		YEVQLGGSMVSMSGCRRKCRKQVVQKACCP
		İ	1			
			l			GYWGSRCHECPGGAETPCNGHGTCLDGMDR
ĺ	ĺ	ĺ	I		ĺ	NGTCVCQENFRGSACQECQDPNRFGPDCQSV
}						CSCVHGVCNHGPRGDGSCLCFAGYTGPHCD
	1					QELPVWQELGFPQNNPRLRKAPNCKCLPG*H
	İ				1	RNGLIATPNPCRP
489	1839	Α	3822	934	669	FFFSEMESRSVTRLECSGAISAHLRLLGSSNSP
'0'	1003	1]		***	ASAS*VAGTIGACHHAQLIFVFLVETGFHHVG
l	}	1	}	ł	}	QDGLDLLNLMIHPPRPPKVLGFQA
400	1840	<u> </u>	3825	79	9748	GCOSCWPAWPRLRRRGPASAGARLGRKAPW
490	1040	A	3023	1 '7	7/40	
	1	1 .				GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA
]]	1		Į	1	ASRPEASGDCRAGRETAMATLEKLMKAFESL
ļ		1		1		KSFQQQQQQQQQQQQQQQQQQQPPPP
	1	Ì	1	i		PPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP
	ĺ		1	· -		GPAVAEEPLHRPKKELSATKKDRVNHCLTIC
1					1	ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA
ł		i	l			ESDVRMVADECLNKVIKALMDSNLPRLQLEL
		1	İ]		YKEIKKNGAPRSLRAALWRFAELAHLVRPQK
		ļ	1		1	CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP
	İ	1			ŀ	KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI
1			1		}	RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG
		l	ļ			LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV
ł	1	l	ł	1		KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL
İ						TLHHTOHODHNVVTGALELLQQLFRTPPPEL
1	1	İ	Ì		ł	LOTLTAVGGIGQLTAAKEESGGRSRSGSIVELI
	1				ļ	
}	}	J	J	1	J	AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS
ļ	1	1	ŀ			ESRSDVSSSALTASVKDEISGELAASSGVSTPG
i		1	1	1		SAGHDIITEQPRSQHTLQADSVDLASCDLTSS
		1				ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG
			1		ļ	TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD
ľ	1		1	<u> </u>		GTDNQYLGLQIGQPQDEDEEATGILPDEASEA
ļ.	l .	į.	1]		FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF
		1				VLRDEATEPGDQENKPCRIKGDIGQSTDDDS
		1		1		APLVHCVRLLSASFLLTGGKNVLVPDRDVRV
1		1	1	1		SVKALALSCVGAAVALHPESFFSKLYKVPLD
		1		1	1	TTEYPEEOYVSDILNYIDHGDPQVRGATAILC
	1	1	i	1		GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL
İ				1		ADCIPLLRKTLKDESSVTCKLACTAVRNCVM
		1	1	1	-	SLCSSSYSELGLOLIIDVLTLRNSSYWLVRTEL
		1	ŀ	Į.		1
]	D.	ļ)	J	1	LETLAEIDFRLVSFLEAKAENLHRGAHHYTGL
		1				LKLQERVLNNVVIHLLGDEDPRVRHVAAASL
		1		1		IRLVPKLFYKCDQGQADPVVAVARDQSSVYL
	1			1		KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD
				i	1	VTMENNLSRVIAAVSHELITSTTRALTFGCCE
		ĺ				ALCLLSTAFPVCIWSLGWHCGVPPLSASDESR
]	j	j]	KSCTVGMATMILTLLSSAWFPLDLSAHODAL
		1	1	1		ILAGNLLAASAPKSLRSSWASEEEANPAATK
			1			OEEVWPALGDRALVPMVEQLFSHLLKVINIC
						AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK
		1	1	1		
	1	1	1	}		GKEKEPGEQASVPLSPKKGSEASAASRQSDTS
			1			GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA
	*1	1	1			NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL
1		1	1	1	1	ATLQDIGKCVEEILGYLKSCFSREPMMATVC
	1	1	[VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA
		1		ļ		QRLGSSSVRPGLYHYCFMAPYTHFTQALADA
1	1	1	1	1		SLRNMVQAEQENDTSGWFDVLQKVSTQLKT
		1				NLTSVTKNRADKNAIHNHIRLFEPLVIKALKQ
L			ــــــــــــــــــــــــــــــــــــــ	<u> </u>	ــــــــــــــــــــــــــــــــــــــ	The same of the sa

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Method	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR KAVTHAIPALQPIVHDLFVLRGTNKADAGKE LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ CHKENEDKWKRLSRQIADIILPMLAKQQMHI DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF VTPNTMASVSTVQL WISGILAILRVLISQSTED IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR ARSMITTHPALVLLWCQILLLVNHTDYRWW AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA AKLGMCNREIVRGGALLFCDYVCQNLHDSE HLTWLIVNHQDLISLSHEPPVQDFISAVHRNS AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL ACRRVEMILAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS PPVSSHPLDGDGHVSLETTSPDKDWYVHLVK SQCWTRSDSALLEGAELVNRIPAEDMNAFM MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA AREVTLARVSGTVQQLPAVHHVFQPELPAEP AAYWSKLNDLFGDAALYQSLPTLARALAQY LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL SWHLHHEQIPLSLDLQAGLDCCCLALQLPGL WSVVSSTEFVTHACSLTYCVHFILEAVAVQPG EQLLSPERRTNTPKAISEEEEEVDPNTQNPK VI TAACEMVAEMVESLQSVLALGHKRNSGVPA FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG WSPRPGGDFGTAFPEIPVEFLQEKEVFKEFTYR INTLGWTSRTQFEETWATLLGVLVTQPLVME QEESPPEEDTERTQINVLAVQAITSLVLSAMT VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK LSIRGIVEQEIQAMVSKRENIATHHLYQAWD PVPSLSPATTGALISHEKLLLQNPERELGSMS YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE ADAPAPSSPPTSPVNSRKHRAGVDHSCSQFL LELYSRWILPSSSARRTHAILGSVURSLLVVS DLFTERNQFELMYVTLTELRRVHPSEDEILAQ YLVPATCKAAAVLGMDKAVAEPVSRLLESTL RSSHLPSRYGALHGVLYVLECDLLDDTTAKQL IPVISDYLLSNLKGIAHCVNIHSQQHVLVMCA TAFYLIENYPLDVGPEFSASIIQMCGVMLSGS EESTPSIIYHCALRGLERLLLSEQLSRLDAESL VKLSVDRVNVHSPIRRAMAALGLMLTCMYT GKEKVSPGGTSDPPPAAPDSESVIVAMERVS VLFDRIKGFPCEARVVARILPQFLDDFFPPQ DIMNKVIGEFLSNQQPYPQFMATVVYKVFQT LHSTGGSSMYRDWVMLSLSNFTTGRAPVAMA TWSLSCFFVSASTSP
					302	SNPPASASRVAGITGVHQHAWLIFVFLVEMEF HHVGQAVLKLLISGDLPVSASQSA
492	1842	A	3836	392	-88	VAPSPMIMPDLYFYRDPEEIEKEE*AAAEK\EE FQSEWTAVV/P/EFTATQSEVADWFKDMQVP
						FQSEWTAVV/PEFTATQSEVADWFKDMQVP SVPIQQFPTEDWST*PTMNDWSATSTAQTTE WVRITTEWP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TPSDMNRAFETDTQSIGEKNRSPSEPDYFERK
			!			KFKRS*EKAHIRYKIDQPEDIPLK\EFLCKHSK CTATLSMRNMSLMKKKCSFSEEF\LAFFPSLL VCHLLAIKLGFYIEIHLTTFNNTF
494	1844	A	3845	2	352	FFFLRRSL/DSVAQAEAQWL\ELGLLQAPPPGF KPISLP\GLPSSWDYGRPPPCPANFCIF/M*RRG FTVLARMVLIS*PCDPPTLASQGTAITGMSYH ARPQDIDFLYAHQGRCWFRLL
495	1845		3847	1774	40	DIFFRAKEGMGQDEAQFSVEMPLTGKAYL WADKYRPKPRFFNRVHTGFEWNKYNQTHY DFDNPPPKIVQGYKFNIFYPDLIDKRSTPEYFL EACADNKDFAILRFHAGPPYEDIAFKIVNREW EYSHRHGFRCQFANGIFQLWFHFKRYRYRR* RPWGTAGRCPRGHSKGASVKLVVTPGPLSGL QGRGFTSHLRPHLSFARPQFPPI*KGGHH*AC HGELRRHWDRLA*GPDATEGALGASFEHEG GQQPPADLTVQADTLHRPSARLGGAHRACPK RRPHRVLWRWARGAWAWRCQAREKQETQG QPCHITGHPLGREAEPAAAGAAPALAHRPPF ARTGSTE\PGPCWRPIRHCRRDPLWTPTLC\nd WPPTHPVLAGGVHFPAAG/IGGCVEVPVSVN VMGTKSH*AVLPPPPSTGPGQGLPEGWGLE KGEGLPPGIPPGLLTGPWSMRPVTPSFAHIR TVAPSHSPFSGQEGRGPHGCHSPGR\SGP\AGR LVLQHPTGTSPTEAKRKVPPGPPEGHPTSPVT SPRPPTAPPRHPASSGNSSVCFSKKTCRWEKK SFVLMELAYWQDRMFF
496	1846	A	3849	830	442	AKSPLPLG*IQWR/NLGSLKLRLPGFK*FTCLG LLSSWDYRSLPPRPVNFCILVELGFHHVDQAG LKLLTSSALPALASQSAEITGMSHRIWPLPLLR RPPVIRIRAPPQRLPFNLITSLKALSPNMATF
497	1847	A	3859	2	393	ALRKTRRDGIARTGAQPAASWKGTNNYPWR LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVPR RFPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ VPIL
498	1848	A	3860	253	634	KNASTVYSSQGDPKSFFFLLRWSLALVAQAG EQ*RDLSSLQPPPPGFK*FSCLSLPSSWD\YRCP LPCLANF*FLVETGFHHVGQADLKLLTSGDP PTSASESAGITGVSHRAWPRIHFLYWKTFFL
499	1849	Α .	3863	423	263	APSQISVAFLYAA/DKLFEKEI*KKIPFIIAS/DKI KIGINLTKEVKYLYTENYITLMKEIK/DTDKW KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP MTFFTEIEKSIIKFIWNHKKPPNTQSNIEQKE*S FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI LVGIALNL
500	1850	A	3865	2	15246	LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP DLRPWASDLDIMGDAEGEDEVQFLRTDDEV VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP TSNAQNVPPDLAICCFVLEQSLSVRALQEML ANTVEAGVESSQGGHRTLLYGHAILLRHAH SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE ACWWTMHPASKQRSEGEKVRVGDDIILVSVS SERYLHLSTASGELQVDASFMQTLWNMNPIC SRCEEGFVTGGHVLRLFHGHMDECLTISPADS DDQRRLVYYEGGAVCTHARSLWRLEPLRIS WSGSHLRWGQPLRVRHVTTGQYLALTEDQG LVVVDASKAHTKATSFCFRISKEKLDVAPKR DVEGMGPPEIKYGESLCFVQHVASGLWLTYA

KLDRLEASSGILEVLYCVLESPEVLNIGENER KSIISLLDKHGRNIK VLDVLCSLCVGGVAR RNODLITENLLPGRELLOTNILNYVTSIRPI IFVGRAEGITQYSKWYFEVMYDEVTPITA ARROWALTEGYTPYFOAGEGWGGNGA GDDLYSY GFDGLHLWTGHVARPVYTSPGGH LAPEDVISCCLIDSYSISTRINGCPVGGVEE NLDGLFFPVSFSAGVKVRFLLGGRHGGFK LAPEDVISCCLIDSYSISTRINGCPVGGVEE NLDGLFFPVSFSAGVKVRFLLGGRHGGFK ROPHLVOPSRCLSHTDFVPCPVDTVQIVLPPI LERIRKLAENIEL WALTRIEGUTYGFVX DWKRLHPCLVDFHSLPEPERNYNLQMSGETI KTLLALGCHVGMADEKAEDNLKKTKLTKT MMSNGYKPAPLDSHYRLITPAQTITLVDRIA NGHNVWARDRVQGWSYSAVQDIPARRIX LVPYRLLDEATKRSNDSLCQAVATILGYG NIEPPOGPSQVENQSRCDRVBIFAGKSYT QSGRWYFEFEAVTTGEMRYGWAPPELRYD GSGRWYFEFEAVTTGEMRYGWAPPELRYD GSGRWYFEFEAVTTGEMRYGWAPPELRYD GSGRWYFFEAVTTGEMRYGWAPPELRYD GSGRWYFFEAVTTGEMRYGWAPPELRYD GSGRWYFFEAVTTGEMRYGWAPPELRYD GSGRWYFFEAVTTGEMRYGWAPPELRYD GSGRWYFFEAVTTGEMRYGWAPPELRYD GSGRWYFFEAVTTGEMRYGWAPPELRYD GSGRWYFFEARDRYD ELGADELAYVPNGHRGGRWHLGSEPFGRY QPGDVVGCMIDLTENTIETLIGHVAMDSG ETAFREIEIGDGFLPYCSLGPGQVGHLNLOG VSSLRFFAICGLQEGFEPFANNQRPVTTWF KGLDQFEPVPLEHPHYEVSRYDGTVDTPPCL LTHKTWGSQNSLVEMLFLRSJRVQFHQHDI CTAGATFLAPPGLQPAEDEAKRAEPDPDYY NLRRSAGGWSEABNGKEGTAKEGAFGFFA AGGGAQPARAFNEKDATTEKNKKRGFLFA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PHILITTTYYSYRVFAQQEGEVAGWVY PDYHOHDMSPDLSKVRVYTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGGLVDLATGLMTTANGRESNTFFQVEFY TKLFPAVFLPTHONVIQFELGKQKDIMPLS, AMFOSERKNPAPCCPPRLEAMMPVSWSI MPNIFLQVETRRAGERLGWAVQCQEPTITM ALHEPERNCKMDFLESSERLDLGFFINTHL YRAVCALGNRVAHALCSHYDQAQLLHAL DAHLPGERAGYDLLISHELDLGFRISHTTLINGFELGKGCRIMPLS, AMFOSERKNPAPCCPPRLEAMMPVSWSI MPNIFLQVETRRAGERLGWAVQCQEPTITENTHL YRAVCALGNRVAHALCSHYDQAQLLHAL DAHLPGERAGYDDLLESSERLDLGFRISHTD KRYSLAGNRVAHALCSHYDQAQLLHAL DAHLPGERAGYDDLLESSERLDLGFRISHTD KRYSLAGNRVAHALCSHYDQAQLLHAL SEYIVPLTETRAITLPPGRSTENGEHPRHGLI GVGVTTSLPPHHFSPPCTVAALPAAGAAR SEYIVPLTETRAITLPPGRSTENGEHPRHGLI GVGVTTSLPPHHFSPPCTVAALPAAGAAR SEYIVPLTETRAITLFPGRSTENGEREED EKEEDEEETAQEKKDEEKEEEAAGGGREED EKEEDEETAGRKABLEKKQRUKKWAGGED DVKQULKMEEPSVLQMCHLLEVFCDQEL HRVESLAAFAERVVDKLQANGRSRYGGLGH KWESLAAFAERVDLALLGCLGGGGGH RDPYGASVEROPTVPLELLECTGGREED		070 70	1/24	OF A	D	Dungling	DATE AND ADDRESS OF THE PARTY O
nucleotide sequence wence unce contemporary with the contemporary							
Deathon Deat			noa				
uence marce 09/496 correspondi ng to fist naid residue nd old residue competide compe			İ				
uence 914 mg to first amino acid residue of peptide residue of peptide sequence residue o	cotide	seq-					
amino acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence per sequence	seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
residue of peptide sequence y=Tyrosine, X=Uaknovn, ==Stop codon,	uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
residue of peptide sequence y=Tyrosine, X=Uaknovn, ==Stop codon,					amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
peptide sequence enciclotide insertion possible sequence pucleotide insertion pucleotide insertion pucleotide insertion pucleotide insertion pucleotide insertion pucleotide insertion pucleotide insertion pucleotide publication publi					residue of		Y=Tyrosine, X=Unknown, *=Stop codon.
mucloside insertion APPEKALRI, GVLKKKAMLHQEGHMDDALS TRCQQEESQAARMIHSTNOLTNO, PIKISLIDSE GKPRGSOPPAGTA LPIEGVILS, LODULIYEEPP EDLQHEEKOSKLRSLRNRQSLFQEEGMLS, MILL LNCIDRLINVYTTAHPAFAGGEAASWKE VINLLYELLASLIRGNRSNCALFSTHLDWL, WILL KLDRLEASSGILEVLYCVLIEEPPVLNIOCEN- KSIISLLDKHGRNHKVLDVLCSLCVCNOVAA- RSNQDLITENLLQREGLLLQTHLNIVYTSIREP FYGRAEGTTQYSKWYFEWM/DEVTPFLTA- ATHLRYGWALTEGPLLIQTHLNIVYTSIREP FYGRAEGTTQYSKWYFEWM/DEVTPSCAG- ATHLRYGWALTEGPLLIQTHLNIVYTSIREP FYGRAEGTTQYSKWYFEWM/DEVTPSCAG- ATHLRYGWALTEGPLLIQTHLNIVYTSIREP FYGRAEGTTQYSKWYFEWM/DEVTPSCAG- ATHLRYGWALTEGPLLIQTHLNIVYTSIREP FYGRAEGTTQYSKWYFEWM/DEVTPSCAG- ATHLRYGWALTEGPLLIQTHLNIVYTSIREP FYGRAEGTTQYSKWYFEWM/DEVTPSCAG- ATHLRYGWALTEGPLLIQTHLNIVYTSIREP FYGRAEGTTQYSKWYFEWM/DEVTPSCAG- ATHLRYGWALTEGPLLIQTHLNIVYTSIREP FYGRAEGTTQYSKWYFEWM/DEVTPSCAG- ATHLRYGWALTEGPLLIQTHLNIVYTSIREP FYGRAEGTTQYSKWYFEWM/DEVTPSCAG- ATHLRYGWALTEGPLLIQTHLNIVYTSIREP FYGRAEGTTQYSKWYFEWM/DEVTPSCAG- NEGPTIC APPELLIC AND ATHLRYGWALTEGRAGET AND ATHLRYGWALTE]				peptide	•	
APPFKALRL GVLKKK AMI-HOEGHMODAL TROOGESQARMISTNOL NOPISLDS GKPRGSGPPAGTAL PIEGVILS LODIL HYEEP ED LOHERC OSKL RENROS LOFEGEMG. SM LINCIDRL NYTT A AHF AEF AGEE ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY VILLY SLLASLIR GRORES ALSWEK VILLY SLLASLIR GRORES ALSWEK VILLY VILLY SLLASLIR GRORES ALSWEK VILLY STATE ALSWEK VILLY S	!				• •		
TRCQGESQAARMIHSTNOLTVOP(ISILIDE) GKPRGSOPPAGTA LPIEOVILS(DOLIVTEPP EDILQHEEKOSKLRSLRNROSLFGEEGMLSM LNCIDRINVYTTA-HPAFFAGEAASWEE VNILLYELLASLIRGNRSNCALFSTHLDWLUS KLDELEASSGILEVLYCVLIESPEVLNIOCEN- KSIISLLDKHGRNHKYLDWLCSLCVCNGVAA RSNQLITENLLPGRELLLQTNLINVYTSIEP IPVGRAEGITQYSKWYPEVMYDEVTPFI.TA ATHERVGWALTEGYTPYPGAGEGWGNGY GDDLYSYGFDGLHLWTGHVARTYSPGGH LAFEDVISICALISTYSPGGHEK THE TORTH TO THE THE TORTH TO THE T					Soquence		
GKPRGSOPPAGTALPIEGVILSLODLITYEEPP EDLQHEEKOSSLIKSRNROSLOFEEGMIS.M LNCIDRLNVYTTAAHFAEFAGEAAESWEE VNLLYSLLASLIRGNRSNCALEPSTILDWILVS KLDRLEASSGILEVLYCVLIESPEVLNIIQENF KSIISLDKHGRNIRKVLDVILCSLCVCNOVAA RSNQDLITENLLDGRELLLQTHLINTYVTSIEPP FVGRAEGTTQYSKWYFEWMVDEVTFPLTAA ATHLRWGWALTEGTYPYPGAGEGWGNGV GDDLYSYGFDGLHLWTGHVARPVTSPGQHFE IFVGRAEGTTQYSKWYFEWMVDEVTFPLTAA ATHLRWGWALTEGTYPYPGAGEGWGNGV GDDLYSYGFDGLHLWTGHVARPVTSPGQHFE LAPEDYISCCLLSYPSISFRINGCPYGGVFES NLDGLFFFVVSFSAGVKVRFLLGGRIGEFKF LPPGYAPCHEAVLFRERHLEFIEKFVREG RGPHLVGPSRCLSHTDFVPCPVDTVQVLPPP LERIREKLAENHELWALTRIEGGWTYPFVR DNRSLHFCLVDFHSLFEFERNYNLQMSGED KTLLALGCHVGMADEKAEDNLKKKLPKT MMSNGYKPAPLDLSHVRLTPAQTTLVTRLA- NGHNVWARDRVGGWSSACDRVTRPRAEKSYTVT GSGRWYFEFEAVTTGERRYGWARPELRFDY ELGADELAYVRNGIRGGRWHLGSEFFGRPY QCGDVYGCMIDLTGERRYGWARPELRFDY ELGADELAYVRNGIRGGRWHLGSEFFGRPY QCGDVYGCMIDLTGERRYGWARPELRFDY ELGADELAYVRNGIRGGRWHLGSEFFGRPY QCGDVYGCMIDLTSHTHITTINGEVLMSDSG ETAFREIEIGDGFLPVCSLGFGQVGHLNLGGG VSLRFFALGGLGGEFFPANNMGRPYTTWF KGLPGPEFPVLEPPHYEVSRVDGTVDTFPCL LTHRTWGSNISLVEMILFLSLEYOPHOTHE CTAGATPLAPPGLQPPAEDEARAEPDDYT NLRRSAGGWSEARMKEGTAKEGAFGTP AGGEAQPARAENEEDATTSKNKKRGFLFKA KKVAMMTQPPATPTLPRHDVVDADNRDI PHILNTTTYYSVRVPAGGPSCVWAGWV PDYNGIDMSFDLSKRWITHAENGEPGRP AGGEAQPARAENEEDATTSKNKKRGFLFKA KKVAMMTQPATPTLPRHDVVDADNRDI PEILINTTTYYSVRVPAGGPSCVWAGWV PDYNGIDMSFDLSKRWITHAENGEPGRP AGGEAQPARAENEEDATTSKNKKRGFLFKA KKVAMMTQPATPTLPRHDVVDADNRDI PEILINTTTYYSVRVPAGGPSCVWAGWV PDYNGIDMSFDLSKRWITHAENGEPGRP AGGEAQPARAENEEDATTSKNKKRGFLFKA KKVAMMTQPATPTLPRHDVVDADNRDI PHILNTTTYYSVRVPAGGPSCVWAGWV PDYNGIDMSFDLSKRWITHAENGEPGRP AGGEAQPARAENEEDATTSKNKKRGFLFKA KKVAMMTQPATPTPRIPHDVADNRDI PHILNTTTYYSVRVPAGGPSCVWAGWV PDYNGIDMSFDLSKRWITHAENGEPGRP AGGEAQPARAENEEDATTSKNKKRGFLFKA KKVAMMTQPATPTPRIPHDVADNRDI PHILNTTTYYSVRVPAGGPSCVWAGWV PDYNGIDMSFDLSKRWITHAENGEPGRP AGGEAQPARENEEDATATSKNKKGFLFKA KROMPATATRIBERFRERPGEPGED PHILNTTYYSVRVPAGGERGRP AGGEAPPARENGERGRP AGGEAPPARENGERGRP AGGEAPPARENGERGRP AGGEAPPARENGERGRP AGGEAPPARENGERGRP AGGEAPPARENGERGRP AGGEAPPARENGERGRP AGGEAPPARENGERGRP AGGEAPPARENGERGRP AGGEAP							
EDLOHEKÇİŞKLRSLENIRQİS FÖÇEĞIMLEN LINCIDELIN VILYELLAŞLIRÖNRƏNCALEŞTINLDIVİŞ VILYELLAŞLIRÖNRƏNCALEŞTINLDIVİŞ KLDELAŞÇILE VILYÜLÜĞÜN KLDELAŞĞILE VILYÜLÜĞÜN KİLDELAŞĞILE VILYÜLÜĞÜN KİLDELAŞĞILE VILYÜLÜĞÜN KİLDELAŞĞILE VILYÜLÜĞÜN KİLDELAŞĞILE VILYÜLÜĞÜN KİLDELAŞĞILE VILYÜLÜNÜLÜN VISIRPI İFÜĞRAĞĞITÜNÜN ÜN VISIRPI İFÜĞRAĞĞITÜNÜN ÜN VISIRPI İFÜĞRAĞĞITÜNÜN ÜN VISIRPI İFÜĞRAĞĞITÜNÜN ÜN VISIRPI İFÜĞRAĞĞITÜNÜN ÜN VISIRPI İFÜĞRAĞĞITÜNÜN ÜN VISIRPI İFÜĞRAĞĞIĞIN ÜN ÇÜN ÇÜN ÇÜN ÇÜN ÇÜN ÇÜN ÇÜN ÇÜN ÇÜN	, ;						
LNCIDELNYTTAAHFAGEEAABSYEL VNLLYELASLIGNRSNCALFSTNLDWLVS KLDRLEASSGILEVLYCVLESPEVLNIQENS KLDRLEASSGILEVLYCVLESPEVLNIQENS KSISLLDKHGRNIKVLDVLCVGVAA RSNODLITENLLPGRELLLQTNLINVTYSEPP IFYGRAEGITQYSKWYFEVMVDEVTYPLTAA ATHLRYGWALTEGYTPYROAGEGWGGNGV GDDLYSYGFDGLHLWTGHVARVTSYGOGY HEAPTOVISCALDLSYPSISTRINGGPVGGVFE NLDGLFFPVSFSAGVKVRFLLGGRHGGFK LPPGYAPCHEAVLPRERLHLEPKEYREGG RGPHLVOPSRCLSHTDFVPCPVDTVQTVLYPVR DNKRLHPCLVDPHSLPBFERNYNLOMSGET ILLALGGHVGMADEKAEDNIKKTLPKT MMSNGYKPAPLDLSHVRLTPAQTILVDRIA NGHNIWARDRYGOGWSYSAVQDIPARRIN LVPYRLLDGATKRSNRDSLCQAVRTLLGG NIEPPDQEPSQVENGSCRDRYRURRABKSYTT QSGRWYFFEAVTTGEMRVGWARPELRPD ELGADELAYVFNGHRAGWHIGSEPFGRAY QFGDVVGCMDLTENTIFITLGSVLMSDSG ETAFREEIGDGFLFVCSLGPGQVGRHLNGCOG VSSLRFFAICGLGGFEFFANNGRVTTWFF KGLRPFPPPLEHHYVSRVDGTVDTPPP LTHTWGSQNSLVEMLFLRS,LPVQFHQHFI CTAGATFLAPPGLQFPAEDARAAEPPDTY KGLRPFPPPLEHHYVSRVDGTVDTPPP NLRRSAGGWSEARNGKEGTAKEGAPGGTFY AGGEAQPARAENSKDATTEKNKKRGFLFK KKVAMMTQPPATTPLRPLHPUVPADNRDK KKVAMMTQPPATTPLRPLHPUVPADNRDK KKVAMMTQPATTPLRPLHPUVPADNRDK KKVAMMTQPATTPLRPLHPUVPADNRDK KKVAMMTQPATTPLRPLHPUVPADNRDK KKVAMMTQPATTPLRRLPHDVPADNRDK KKEDEREGEGEGED DVKQLYBLERGEGGGEGED DVKQLYBLERGEGGGEGED DVKQLYBLERGEGGGEGGED DVKQLYBLERGEGGGEGGED DVKQLYBLERGEGGGGTEGGDE DVKQLYBLERGEGGGGTEGGDE DVKQLYBLERGEGGGGTEGGDE DVKQLXMEEPSVLLQWCHLLEVFCDQGL HRVESLAAFAERYVDKLQANQRSRYGLLK FSMMAAETARRTRFRPPQCGJONMLLGFGG DVKQLXMEEPSVLLQWCHLLEVFCDQGL LEGLLOMKLPSVVLLQWCHLLEVFCDQGL UGSELVRAMFSLLECLGQGGGGGGGGGGGGGGG							
WILLYELLASILIRONRSINCALFSTINLDUY, KLDRLEASSGILEVILYCVLESPVLYNIOEN KSIISLLDKHGRNHK VLDVLCSLCVCNGVAA RSNODLITENILLPGRELLLOTNLINYVTSIRPI FFGRAEGTTQYSKWYFEWYDDVTPFTLTA ATHLRVGWALTEGYTYYFAGEGWGGNG GDDLYSYGFDGLHLWTGHARVYTSPGQH LAFEDVISCCLDLSVPSISFRINGCPVGGVFE NLDGJFFFVVSSFAGVKVRFLIGGRHGEFK LPPGYAPCHEAVLPRERLHLEPIKEYREG GDDLYSYGFDGLHLWTGHARVYTSPGQH LAFEDVISCCLDLSVPSISFRINGCPVGGVFE NLDGJFFFVVSSFAGVKVRFLIGGRHGEFK LPPGYAPCHEAVLPRERLHLEPIKEYREG RGPHLVGPRGCLSHTDFVPCPVDTVQIVLPPI LERIBEKLAENIFELWALTRIEGGWTYGPYW DMKRLHPCLVDFHSLPPEERNYLOMSGET XTLLALGCHVGMADEKAEDNILKKTKLPKT MMSNGYKPAPLDLSHVRLTPAQTTLVTDL NGHNVWARDRVGGWSYSAVQDIPARRAF LVPYRLLDBATKRSNRDSLCQAVRTLLGYGY NIEPPDQEPSQVENGSKCDRWIFRAEKSYTV QSGRWYFEFEAVTTGEMRVGWARPELRPD ELGADELAYVPNGHRQGWHINLGQ VSGRWYFEFEAVTTGEMRVGWARPELRPD ELGADELAYVPNGHRQGWHINLGQ VSSLRFFALCGLGEGEEPPANNGRPVTTUPF KGLPGEPPVLEHPHYVSKYDGTVDTPDPC LTHETWGGONSLVEMIFLRLSLPVQFHQHE CTAGATPLAPPGLPPADEARAAEDPDPY NLRRSAGGWSEAENGKEGTAKEGAPGGTPP AGGEAQPARAENEKDATTERNKKRGHFLKF KKVAMMTQPPATPLPRLPHUVYPADNRDI PEILINTTTYYYSVRVFAGQEPSCVWARWY PDYNGHMSMPLOKKWRGTLFKK KKVAMMTQPATPLPRLPHUVYPADNRDI PEILINTTTYYYSVRVFAGQEPSCVWARWY PDYNGHMSMPLOKKWRGTLFKK KKVAMMTQPATPLPRLPHUVYPADNRDI PEILINTTTYYYSVRVFAGQEPSCVWARWY PDYNGHMSMPLOKKWRGTLFKK KKVAMMTQPATPLPRLPHUVYPADNRDI PEILINTTTYYYSVRVFAGQERSCVWARWY PDYNGHMSMPLOKKWRGTLFKK KKVAMMTQPATPLPRLPHUVYPADNRDI PEILINTTTYYYSVRVFAGQERSCVWARWY PDYNGHMSMPLOKKWRGTLFK KKVAMMTQPATPLPRLPHUVYPADNRDI PEILINTTTYYYSVRVFAGQERSCVWARWY PDYNGHMSMPLOVETRAGERLGWAVQCQCPHTMM ALHIPEFNCKOMDLELSERLDLQRHSHTILL YRAVCALGNRVAHALCSHVQAQLLHAL DHHPGPRACMDPLELSERLDLQRHSHTILL YRAVCALGNRVAHALCSHVQAQLLHAL DHHPGPRACMDPLELSERLDLQRHSHTILL YRAVCALGNRVAHALCSHVQAQLLHAL DHHPGPRACMDPLELSERLDLQRHSHGLI GVGVTTSLRPPHHISPPCVAALPAAGAAEA ARLSPABLELALGRUSVLJOKKKE EEGEEFETTLGSRLMSLLEKVQRLOKKE EKEDEEFTAGRKAGGGH RDPVGASVEPQTVPLKLVSTLLVMGIPGG DVKQULKMIEPSVALQMCHLLEYFCDQE LEGELPEEETTLGSRLMSLLEKCHQRICHURGHG GEEPPEETTLGSRLMSLLEKCHQULCHQMGQ VGSELVRAMFSLLHCQGGGGGGELLBALPR YTISPSSVEDIMGELLECLGGGRICHQMG							
KLDRLEASSGILEVLYCVLESPEVLNIGENER KSIISLLDKHGRNIK VLDVLCSLCVGGVAR RNODLITENLLPGRELLLQTNLINYVTSIRPI IFVGRAEGTTQYSKWYFEVMYDEVTPITA ATHLRVGWALTEGYTPYGAGEGWGGNGV GDDLYSYGFDGLHLWTGHVARPVYTSPGGH LAPEDVISCCLIDSYSISTRINGCPVGGVFES NLDGLFFPVVSFSAGVKVRFLLGGRHGGFK LAPEDVISCCLIDSYSISTRINGCPVGGVFES NLDGLFFPVVSFSAGVKVRFLLGGRHGGFK ROPHLVOPSRCLSHTDFVPCPVDTVQIVLPPI LERIRKLAENIEL WALTRIEGUTYGPVX DNKRLHPCLVDFHSLPEPERNYNLQMSGETI KTLLALGCHVGMADEKAEDNLKKTKLTKT MMSNGYKPAPLDLSHYRLTPAQTITLVDRIA NGHNVWARDRVGQGWYSSAVQDIPARRIX LVPYRLLDEATKRSNDSLCQAVRTLLGYG NIEPPDQEPSQVENQSRCDRVBIRALKSYT QSGRWYFEFEAVTTGGMRVGWARPELRPD GRADELAYVFNGHRGGRWHLGSEFFGRPY QFGDVVGCMIDLTENTIETLINGVMARPELRPD ELGADELAYVFNGHRGGRWHLGSEFFGRPY QFGDVVGCMIDLTENTIETLINGVMARPELRPD ELGADELAYVFNGHRGGRWHLGSEFFGRPY QFGDVVGCMIDLTENTIETLINGVMARPELRPD ELGADELAYVFNGHRGGRWHLGSEFFGRPY QFGDVVGCMIDLTENTIETLINGVMARPELRPD ELGADELAYVFNGHRGGRWHLGSEFFGRPY QFGDVVGCMIDLTENTIETLINGVMARPELRPD ELGADELAYVFNGHRGGRWHLGSEFFGRPY QFGDVVGCMIDLTENTIETLINGVMARPELRPD ELGADELAYVFNGHRGGRWHLGSEFFGRPY QFGDVVGCMIDLTENTIETLINGVMARPELRPD FGLIGHTYTYYSYSVRYAGQEFFANNGRPYTTWP KGLPGFPPPLEHPHYEVSRNDGTVDTPPEL LTHRTWGSQNSLVEMILFLRSLPVOFHOHPI CTAGATFLAPPGLOPPAEDEARAEPDPDYT NLRRSAGGWSEABNGREGTAKEGAFFLFAK KKVAMMTQPPATPTLPRLPHDVVPADNRDD PHILINTTTYYSVRYAGQEGFEYMAGWY PDYHGHDMSPDLSKVRYVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTTTANGKESNTFFQVEP TKLFPAVFVLPTHONVIQFELGKQKDIMPLS. AMFGSERKNPAPCCPPRLEAMMPVSWSI MPNIFLQVETRRAGERLGWAVQCQEPLITM ALHPEPRRCMDHLELSBELLDGRPRSHTHL YRAVCALGNRVAHALCSHVDQAGLHALA LDAHLPGFLRAGYDLLISHELGKRGRRSMM SEYIVPLTPETRAITLPPGRSTENGFHPRHGLI GVGYTTSLPPHHFSPPCTVAALPAAGAAAE ARLSPAILEALARDKALRMLGSRYBGGGH RDPYGASVEPQFVPVLKLVSTLLVMGHEGE EKEEDEEETAQEKEEEEAAEGGREEDE EKEEDEEETAQEKEEEEAAEGGREEDE EKEEDEEETAQEKEEEEAAEGGREEDE EKEEDEEETAGERRAMLLEKVRUNKKAGEGI HVESSLAAFAERVVDKLQANQRSRYGLLIKL FSMTAASTAAFLERVGRPVEVELEWEEDE							LNCIDRLNVYTTAAHFAEFAGEEAAESWKEI
KSIISLLDKHGRNHKVLDVLCSLCVCKGVAN RSNQDLITEL LPGRELLLQTNLLNYTSIRP IFVGRAEGTTQYSKWYFEVMYDEVTPLTA ATHLRVGWALTEGYTPYPGAGEGWGGNOG GDDLYSYGFDGLHLWTGHVARPVTSPGQH LAFEDVISCCLDLSVPSISFRINGCPVGGVES NLDGLFFPVYSFSAGVKVFRLGGGRGEFKF IPPGYAPCHEAVLFRER.HLEPIKSPREGG ROPHLVGPSRCLSHTDFYCEVDTVQVLPIP LERIREKLAENIHEL WALTRIEGGWYTGPVR DNKRLHPCLVFPSLSPERNYNLOMSGET KTLALGGHVGMADEKAEDNLKKTKLPKT MMSNGYKPAPLDLSHVRLATQTTLVDRIA. NGHNVWARDRVGGGWSYSAVQDIPARRNF LVPYRLLDBATKRSNRDSLQAVARTLLGVG NIEPPDQEPSQVENQSRCDRVRIFRAEKSYTV QSGRWYFEFAAVTTGGRWHLGSEPFGRPV QGDVVGCMIDLTENTIFTLNGEVLMSDSG ETAFREEIGDFLPVCSLGGVGHLNLGGV VSSLRFALCGLQEGFEPFAINMQRRYTTWF KGLOFFEPVLEPHYVEVSRVDGTVTDFPCL LTHRTWGSONSLVEMLFLRLSLPVOPHOHF CTAGATPLAPPGLOPPAGEAARAPPDPDYS NLRSASGGWSEAENGKEGTAKEGARGFFFA KKVAMMTQPPATPTLPRLATDVYPADNDL PEILINTTTYYSVRVEAGGEPSCVMAGWV PPDYHOHDMSFDLSKVRVVTVTMGDEGGRW VSIGLSVAGVAGWTPPHHONVOFFELGKOKNIMPLS AKPOSERNEN SKLCSNCYMVWGGDPVSPQQGGRSHTD VIGCLVDLATGLMFTANGKESNTFFQVEPP TLLFPAVFVLPTHQNVTQFELGKOKNIMPLS AMPGSERKNPAPQCPPRLEMQMLMPSWSS MPNHFLQVETRRAGERLGWAVQCOEPLTM ALHPEENRGMULLSERLGWAMGNESTFFQVEPP TLLFPAVFVLPTHQNVTQFELGKOKNIMPLS AMPGSERKNPAPQCPPTLEMQMLMPYSWS MPNHFLQVETRRAGERLGWAVQCOEPLTM ALHPEENRGMULLSERLGWAMGNESTFFQVEPP TLLFPAVFVLPTHQNVTQFELGKOKNIMPLS AMPGSERKNPAPQCPPPLEMQMLMPYSWS MPNHFLQVETRRAGERLGWAVQCOEPLTM ALHPEENRGMULLSERLGWAMGNESTHGLWEPP TLLFPAVFVLPTHQNVTQFELGKOKNIMPLS AMPGSERKNPAPQCPPPLEMQMLMPYSWS MPNHFLQVETRRAGERLGWAVQCOEPLTM ALHPEENRGMULLSERLGWAMGCRSTHGUEPH GGGLVDLATGLAMFTANLARCSNTFFQVEPP TLLFPAVFVLPTHQNVTQFELGKOKNIMPLS AMPGSERKNPAPQCPPPLEMQMLMPYSWS MPNHFLQVETRRAGERLGWAVQCOEPLTM ALHPEENRGMULLSERLGWAMGCRSTHGUEPH GGGLAMFVALALCSHVDQAGALA ARLSPAPLEALARACALRAGLARLGRAUGGRSTHGUEPHGLL HVESLAAFABRYVDKLQANGRSPCGLLK FSMTAAETARTREFSSPCFOVALPAAGAAEA ARLSPAPLEALARKALRIKLERSPCFVAALPAAGAAEA FREEGRAEESKERGLELLEFYCDQDL HVESLAAFABRYVDKLQANGRSRYGLLK FSMTAAETARTREFSSPCFOVALPAAGAAEA FREEGRAEESKFRSLQELLVSKKKEE FFEERSAEESKRSLQELVSKKKEE FFEERSAEESKRSLGELVSHMVVRWAGED VQSPELVRAMFSLLHRQYDGLGELLARAPR.							VNLLYELLASLIRGNRSNCALFSTNLDWLVS
KSIISLLDKHGRNHKVLDVLSLCVCKOVAY RSNODLITENLLPGRELLLQTHLNYYTSIRP FYGRAEGTTQYSKWYFEVMYDEVTPILTA ATHLRVGWALTEGYTPYPOAGEGWGGNGY GDDLYSYGFDGLHLWTGHVARPVTSFGQH LAFEDVISCCLDLSYFSISFRINGCPVGGYFE NLDGLFFFVVSFSAGVEVRFLLGGRHGEFK LPPGVYAPCHEAVLPRERLHLEPKEYFEGGH LAFEDVISCCLDLSYFSISFRINGCPVGGYFE NLDGLFFFVVSFSAGVEVRFLLGGRHGEFK LPPGVYAPCHEAVLPRERLHLEPKEYFEGGH LAFEDVISCCLDLSYFSISFRINGCPVGGYFE NLDGLFFFVVSFSAGVEVRFLLGGRHGEFK LPPGVYAPCHEAVLPRERLHLEPKEYFEGGH LAFEDVISCCLDLSYFSISFRINGCPVGGYFE NLDGLFFFVVSFSAGVEVRFLLGGGHGEFK LPPGVYAPCHEAVLPREGGHTGFTVALOMSGET LPPGVAPCHEAVLPREGGHTGAMSGET LERIBEKLAENHELWALTRIEGGWTYGPVFP LERIBEKLAENHELWALTRIEGGWTYGPVFPP LERIBEKLAENHELWALTRIEGGWTYGPVFPP LERIBEKLAENHELWALTRIEGGWTYGPVFPP LERIBEKLAENHELWATTLOFFT MMSNGYKPAPLOLSHVALTLGVG NIEPPDQEPSQVENQSCCDRVRIFAAEKSYTV QSGRWYFEFEAVTICHBRVGWWARPPLRFD LYPYRLDEATKENSDELCGBRFFGRPV QSGRWYFEFEAVTICHBRVGWWARPPLRFD ELGADELAYVFNGHAGGWWALGEFFGRPV QSGRWYFEFEAVTICHBRVGWWARPPLRFD ELGADELAYVFNGHAGGWHCHNLGG VSSLRFFALGCLGEGFEFYANNGRPVTTWFF KGLPGFPVPLEHPHYEVSKYDOTVDTPPCL LTHKTWGSONSLVEMIFLRLSIPVOPHOPHOPHOPHOLOPHOLOPHOLOPHOLOPHOLOPHO							KLDRLEASSGILEVLYCVLIESPEVLNIIQENHI
RSNODLITENLLPGRELLQTNLNNYTSIRP IFVGRAGGTOYSKWYPEVMYDEVTPITA ATHLRVGWALTEGYTPYPGAGEGWGGNGV GDDLYSYGFDOLHLWTGHVARPYTSPGQHI LAPEDVISCCLDLSVPSISRINGCPVQGVFER ILDGLFFVVSFSAGVKVRFLLGGREFKI LPPFGYAPCHEAVLPRERLHLEPIKEYREGE ROPHLVGPSRCLSHTDFVPCPVDTVQTVPP LERIREKLABNHELWALTRIEGGWTYGPVR DNKRLHPCLVOFHSLPEPERNYNLGMSGET KTLLALGCHVGMADEKA BENNELWALTRIEGGWTYGPVR DNKRLHPCLVOFHSLPEPERNYNLGMSGET KTLLALGCHVGMADEKA BENNELKTLEPKT MMSNGYKPAPLDLSHVRLTPAQITLVDRLA NGHNVWARDRVGGWSYSAVQDIPARRNF LVPYRLLDEATKRSNRDSLOVARTILGGY NIEPPDQEPSQVENGSRCDRVRIFRAEKSYTV OSGRWYFEFEAVTTGEMKOAPRELEPFORPY ELGADELAYVFNGHRGGRWHLGSEFFGRPY OPGDVVGCMIDLIENTIFTLNGEVLMSDSG ETARREIEIGDGFLPVCSLGPGQVGHLN.GQI VSSLRFALGGLGGCFFFANGRPVTTWFF KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCL LTHRTWGSONSLYBMLFELLSLFVCPGHGHF CTAGATFLAPGLQFGAFEARAEPPDTVY NLRRSAGGWSEAENGKEGTAKEAGAPGTPY AGGEAQPARAENEKDATTENKKRGFLFK, KKVAMMTQPPATPTLPRLHDVVPADNRDI PEIILNTTTYYSVRVFAGGPSCVWAGWV PDYHOHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCVAMVWGGDFSCWAGWV PDYHOHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCVAMVWGGDFSCWAGWS MPNHFLQVETRAGGREGGMAQCQCPLTM ALHPEENREGMULLSERIDLORFHSHTLRI VRAVCALGMNRVAHALCSHVDOAQLHAL DAHLGPLRAGAYYDLLISHLESACRSRRSM MPNHFLQVETRAGGRIGGAAPAGAAAA ARLSPAPLEALRNALGFAVRDGGH ROPYGASVERGPFVVLKLWSTLLWMGHGDE DVKQLKMIBPVTTFEEEEDEEEEGDEEEGEEEDE EKEEDGEETAQEKEDEKEEEAAAGGKEEC LEEGLLOMKLFESVLLGARGRIGGEGEEGEEEDE EKEEDGLEETAGEKEERGERAAGGKEEC LEEGLLOMKLFESVLLGWAGNAGAABA ARLSPAPLEALRNALGFAVRDGGHR ROPYGASVERGFFVVLKLWSTLLWMGHGDE DVKQLKMIBPVTTFEEEEDEEEEGEEEEGEEEDE EKEEDGLEETAGREKEECLAFGRKEEC LEEGLLOMKLFESVLLQMGRABAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA							
IFVGRAEGTTQYSKWYFEVMYDEVTPILTA ATHLRVGWALTEGYTPYGGGEWGGNOG GDDLYSYGFDGLHLWTGHVARPVTSPGQH LAFEDVISCCLDLSVPSISFRINGCPVGGVES NLDGLFFVVYSFSAGVKVKPLLGGRIGEFKF LPPGYAPCHEAVLPRERLHLEPIKEYREG ROPHLVGPSRCLSHITDFYCPVDTVQVLPIP LERIREKLAENHELWALTRIEGGWYTGPVR DMKRLHPCLVOFPHSLPPERPNYNLOMSGGT KTLLALGCHVGMADEKAEDNILKKTKLPKT MMSNGYKPAPLDLSHVRLTAQTITLOPICA NGHNWARDRVGGGWSYSAVQDIPARRNIP LVPYRLDEATKRSNRDSLCQAVRTILGYG NIEPPDQEPSQVENQSRCDRVBIFRAEKSYTV QSGRWYFEFEAVTTGEMRVGWARPELRPD ELGADELAYVRNGHRGGWAVFULGSPFGRPV QCDVVGCMIDLTENTIFITLNGEVLMSDSG ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQI VSSLRFFALCGLGEGFEPAMMQRVTTWFF KGLPQFEPVPLEHPHYEVSRVDGTVQTPPCL LTHRTWGSDNSLVEMLFLRSLPVQFHQHFP CTAGATPLAPPGLQPPAEDEARAAEPPDDYT MKRSAGGWSAENOKEGTARGAPGGTP AGGEAQPARAENEKDATTEKNKKRGFLFAA KKVAMMTQPPATPTLPRPHDVVPADNRDI PEILINTITYYYSVRVFAQGPSCVWAGWV PPYHQHDMSFDLSKVRVVTVTMGDEGGNV HSSLKCSNCYMVWGGDPYSPQQQGRISHTD VIGCLVDLATGLMTTPANGKESNTFPQVEP TLLPPAVFVLPTHQNNTQELGGKFPFOVEP TLLPPAVFVLPTHQNNTQELGGKFPFOVEP TLLPPAVFVLPTHQNNTQELGGKFFNHLKL YRAVCALGNNRVAHALCSHVOQQLLHAL DAHLPPENRCMDLELSERLDLGRFHSHTLKI YRAVCALGNNRVAHALCSHVOQQLLHAL DAHLPPELRAGYYDLLISHLESACRSRRSM SEYIVPLTFETRAITLFPPGRSTENGHRGGL GVGYTTSLRPAHFSPCTVALLAPAGAAGA ALSPAIPLEARDVALLSHLESACRSRSM SEYIVPLTFETRAITLFPPGRSTENGHRGGL GVGYTTSLRPHFSPPCTVALLAPAGAAGA ALSPAIPLEARDVALLSHLESACRSRSM SEYIVPLTFETRAITLFPRGSTENGHRGGL GVGYTTSLRPHFSPPCTVALPAAAAGAAGA ALSPAIPLEARDVALLSHLESACRSRSM SEYIVPLTFETRAITLFPRGSTENGHRGGL GVGYTTSLRPHFSPPCTVALPAAAAGAAGA ALSPAIPLEARDVALLSHLESACRSRSM SEYIVPLTFETRAITLFPRGSTENGHRGGL GVGYTTSLRPHFSPPCTVALPAAAAAGAAGA ARLSPAIPLEARDKALRMLGEAVRDGGOH RDPVGASVEPCFVPVLLVSTLLVMGIFGGE EKEEDEEETAQEKEEGEBEEEE EKEEDEEETAGEREECE LEEGLLQMKLPESVKLQMCLLEHTPCDQL HWESLAAPAGRYDLLJSHLAPRACHAGA ARLSPAIPLEARDKALRMLGEAVRDGGOH RDPVGASVEPCFVPVLLVSTLLVMGIFGGE EKEEDEEETTLOSKLMSLLEKVRLVKKKEE PEERSAEESKRRSLQELVSHMVVRWAGED UKGELAVARSSLLHRQYDGLGGRISLLIVQMGGP VGSPELVRAMFSLLHRQYDGLGGLKLIVQMGGOH VGSPELVRAMFSLLHRQYDGLGGRISLLIVQMGGOH VGSPELVRAMFSLLHRQYDGLGGRISLIVQMGGOH VGSPELVRAMFSLLHRQYDGLGGLIRAPRA							
ATHLRVGWALTEGYTPYPGAGEGWGGNOV GODDLYSYGPGLH.WTGHAVTYSPOGH LAPEDVISCCLDLSVPSISFRNGCPVQGVFE NLDGLFFFVVSFSAGVKVRFLLGGRHGEFKI LPPFGYAPCHEAVLFRER.HLEPJKEYREGE ROPHLVGPSRCLSHTDFVPCPVDTVQVLPPI LERIREKLAENHELWALTRIEQGWTYGPVR DNKRLHPCLVDFHSLEPFERNYNLOMSGET KTLLALGGHVGMADEKAEDNLKKTKLPKT- MMSNGYKPAPLDLSHVRLTPAQTTLVDRLA NGHNVWARDRVQGQWSYAQDIPARRINF LVPYRLLDEATKRSNRDSLCQAVRTLLGYG NIEPPDQEPSQVENQSRCDRVRIFAEKSYTV QSGRWYTEFEAVTTGEMKVAPDIPARRINF LVPYRLLDEATKRSNRDSLCQAVRTLLGYG NIEPPDQEPSQVENQSRCDRVRIFAEKSYTV QSGRWYTEFEAVTTGEMKVARPELREDY ELGADELAYVYNGHRGQRWHLGSEPFGRPV QPGDVVGGCMBLTENTITTLNGEVLMSDSG ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQ VSSRLFFALGGLQGCFEPFAINMQRRYTTWF KGLPGFEPVLEHPHVEVSRVDGTVDTPPCL LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFI CTAGATPLAPPGLQPFAEDEARAAEPPDPTY NLRSASGGWSEAENGKEGTAKEGAPGGTPC AGGEAQPARAENEKDATTELPVQFHQHFI CTAGATPLAPPGLQPFAEDEARAAEPPDPTY NLRSASGGWSEAENGKEGTAKEGAPGGTPC AGGEAQPARAENEKDATTELPVQFHQHFI CTAGATPLAPPGLQPFAEDEARAAEPPDPTY NLRSASGGWSEAENGKEGTAKEGGFFKX KVAMMTQPPATPTLPRLPHDVVPADNRDL FEIILNTTTYYYSVRVFAGGPSCVWAGWV PDYHQIDMSFDLSKVRVVTVTMGDEPGONV HSSLKCSNCVMVWGGDPSYPQQQGRISHTD VIGCLVDLATGLMTTANGKESNTFPQVEPP VKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPFLEMQMLMPYSWSI MNHHLQVETRRAGERLGWAVQCGCPLTM ALHPEENRGMLELSERLDLQRFHSHTLRI YAAVCALGNRVAHALCSHVDQAQLHAL DAHLGPLRAGYYDLLISHLBSACRSRRSM SSYLVPLTFETRAITLFPPGRSTENGIPPRGIGL GVQTTSLRAGYPVLLISHLBSACRSRRSM SSYLVPLTFETRAITLFPPGRSTENGIPPRGIGL GVQTTSLRAGYYDLLISHLESACRSRRSM SSYLVPLTFETRAITLFPPGRSTENGIPPRGIGL GVQTTSLRAGYYDLLISHLESACRSRRSM SSYLVPLTFETRAITLFPPGRSTENGIPPRGIGL GVQTTSLRAGYYDLLISHLESACRSRRSM SSYLVPLTFETRAITLFPPGRSTENGIPPRGIGL GVQTTSLRAGYYDLLISHLESACRSRRSM SSYLVPLTFETRAITLFPPGRSTENGIPPRGIGL GVQTTSLRAGYYDLLISHLESACRSRRSM SSYLVPLTFETRAITLFPFGCDECEGEEDE EKEEDEEETAGEKEEC LEEGLLQMKLPESVKLQMCARABA ARLSPAIPLEALRAKALFBAYDALGALEARA ARLSPAIPLEALRAKALFBAYDALGALEARA FREFEREIGEBEEETEGEEEGEEEDE EKEEDEETTAGREERGERGEGGEED EKEEDEETTARGREFTSPORTHATHATARAESTTERFFT TTATTLFTTATTLFT TTATTLFT TTATTLFT TTATTLFT TTATTLFT TTATTLFT TTATTLFT TTATTLFT TTATTLFT TTATTLFT TTATTLFT TTATTLFT TTATTLFT TTATTLF	ł						
GDDLYSYGFDGLHLWTGHVARPYTSPGQH LAPEDWISCLDLSVPSRAGVKVRFLLGGHGEFK NLDGLFFPVVSFSAGVKVRFLLGGHGEFK NLDGLFFPVVSFSAGVKVRFLLGGHGEFK NLDGLFFPVSFSAGVKVRFLLGGHGEFK LPPPGYAPCHEAVLPRERLHLEPIKEYREGG RGFHLVGPSRCLSHIDFYPCPVDTVQIVLPPI LERIREKLAENHELWALTRIEGGWTYGPVR DNRKHHFCLVDFHSLFPERNYNLQMSGETI KTLLALGCHVGMADEKAEDNIKKTKLPKT- MMSNGYKPAPLDLSHVLPAQTITLORIA NGENVWARDRVGQGWSYSAVQDIPARRNF LVPYRLLDEATKRSNRDSLCQAVRTLLGYG NEEPPDGEPSQVENQSKCDRVRIFRAEKSYTV QSGRWYFEFAVYTGGMRVGWARPELRPD ELGADELAYYFNOHRGGRWHLGSEPFGRPV QPGDVVGCMDLIENTHIFTLNGEVLMSDSG ETAFREIEIGDGFLPVSGGOVGHLNLGQI VSSLRFFAICGLQEGPEPPAINMQRPVTTWF KGLPPGFPVPLEHPHYEVSRVDGTVDTPFCL LTHRTWGSQNSLVEMHLRLSLPVQFHOHFI CTAGATPLAPPGLQPPAEDEARAEPPDTVY NIRRSAGGWSEAENGKEGTAKEGAPGGTP AGGAQPARAENEKDATTERKKRGFLFK KKVAMMTQPATPTLPHDVVPADNRA KKVAMMTQPATPTLPHPHDVVPADNRA KKVAMMTQPATPTLPHPHDVVPADNRA KVAMMTQPATPTLPHPHDVVPADNRA KVAMMTQPATPTLPHPHDVVPADNRA KVAMMTQPATPTLPHPHDVVPADNRA KVAMMTQPATPTLPHPHDVPADNRA KVAMMTQPATPTLPHPHDVPADNRA KVAMMTQPATPTLPHPHDVPADNRA KVAMMTQPATPTLPHPHDVPADNRA KVAMMTQPATPTLPHPHDVPADNRA KVAMMTQPATPTLPHPHDVPADNRA KVAMMTQPATPTLPHPHDVPADNRA KVAMMTQPATPTLPHPHDVPADNRA KVAMMTQPATPTLPHPHDVPADNRA KVAMMTQPATPTLPHPHDVPADNRA KVAMMTQPATPLFLPHDVPADNRA KVAMMTQPATPTLPHPHTDVPADNRA KVAMMTQPATPTLPHPHTDVPADNRA KVAMMTQPATPTLPHPHTDVPADNRA KVAMMTQPATPLFLPHDVPADNRA KVAMMTQPATPTLPHPHTDVPADNRA KVAMMTQPATPLFLPHDVPADNRA KVAMMTQPATPLFLPHDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA KVAMMTQPATPLFLPHTDVPADNRA NTRAETCAR KREDGERSENGENGENGENGENGENGENGENGENGENGENGENGENGE							
LAPEDVISCLDLSVPSISFRINGCPVQGVFES NLDGLFFPVVSFSAGVFLLGGREFKS LPPPQYAPCHEAVLPRERLHLEPIKEYRREGI RGPHLVGPSRCLSHTDFVPCYVDTVQTVLPPI LERIREKLAENIHEL WALTRIEQGWTYGPVR DNRRLHPCLVDFHSLPEPERNYNLQMSGETI KTLLALGCHVGMADEKAEDNLKKTKLPKT MMSNGYKPAPLDLSHVRLTPAQITILVDRLA NGHNVWARDRVGQWSYSAVQDIPARRIA NGHNVWARDRVGQWSYSAVQDIPARRIA NGHNVWAFDRVGGWSYSAVQDIPARRIA NGHNVWAFDRVGGWSYSAVQDIPARRIA NGHNVWAFEFAVTTGEMRVGWAFPELRPD GSGRWYFEFAVTTGEMRVGWAFPELRPD ELGADELAYYFNGHRGGRWHLGSEPFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV QPODVVGCMDLTENTHLNGEVLMSDEFFGRPV GREFTLAFTENTHLNGEVLMSDEFFGRPV GREFTLAFTENTHLNGEVLMSDEFFGRPV GREFTLAFTENTHLNGEVLMSTEFFGRPVEP TRASAGGWSEAENGKEGTAKEGAPGGTPC AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPFATPLTLPHDVVPADNRDL PEILINTTTYYSVRVFAGQEPSCWAGWVY PPYHQHDMSPDLSKVRVTVTTMGDEGGNV PPYHQHDMSPDLSKVRVTVTTMGDEGGNV PPYHQHDMSPDLSKVRVTVTTMGDEGGNV PSYHQLDATGLMTFANKESNTFFQVEP TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS AMFQSBRKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVTRTARGERAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALGWAVQCCPPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALGWAVQCCPPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALGWAVQCCPPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI GROVTTSLRPPHHTSPFCTVAALPAAGAAEA ARLSFAIPLEALROKALRMLGEAVAUCKHE GEGLEPPEEEDEEEEGEEEDEE EKEDEETAQEKEEEDEEEEGEEEDE EKEDEETAQEKEEEDEEEEGEEEDE EKEDEETTLAGKLMGRYGLILHQLAHCGGGG DVKQTLKMIPPSVTTEEEEEDEEEEEGEEEDE EKEDEETTLAGKLMGRYGLILHQLAHCGGGG DVKQTLKMIPPSVTLAKYGRYGLLKAVEKKEE EEGELQMLLMFRSVLLVSTLLVMGFGGD DVKQOLKMIPPSVTLLVSTLLVMGFGGD DVKQOLKMIPPSVTLLVSTLLVMGFGGD DVKQOLKMIPPSVTLLVSTLLVMGFGGD DVKQOLKMIPPSVTLLVSTLLVMGFGGD DVKQOLKMIPPSVTLLVSTLLVMGFGGD TDEEDCPLPEEETTLOSRLMLLEKVLLVKKKEE EFEERSAEESKPSLQELVSHMVVWAQGED VQSPELVRAMFSL	†						
NLDGLFFPVSFSAGVKVRFLLGRHGEFSE LPPPGYAPCHEAVIPREIL HEPIEYPREGI RGPHLVGPSRCLSHTDFVPCPVDTVQIVLPPI LERIREKLAENIHEL WALTRIEGGWTYGPVR DNRCHIPC.VDFHSLFPERNYNLQMSGETI KTLLALGCHVGMADEKAEDNIKKTKLPKT. MMSNGYKPAPLD.SHVLPAQTITLVDRLA NGHNVWARDRVGQGWSYSAVQDIPARRNF LVPYRLDEATKRSNRDSLCQAVRTLLGYG NIEPPDGEPSQVENQSKCDRVRIFRAEKSYTV QSGRWYFEFAVTTGEMRVGWARFELRED ELGADELAYVFNOHRGGWHLGSEPFGRPV QPGDVVGCMIDLTENTIIFTLNGEVLMSDSG ETAFREIEIGDGFLPVSLGPGQVGHLNLGG) VSSLRFFAGCLGGEFFAINMQRPVTTWF KGLPQFEPVPLEHPHYEVSRVDGHLNLGG) LTHRTWGSONSLVEMLFLRLSLPVQFHOHFI CTAGATPLAPPGLQPPAEDEARAAEPDPDYF NIRRSAGGWSEAENGKEGTAKEAGPGGTTV AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVPADNRDI PEILINITTYYSVRVPAGGEPSCVWAGWV PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSFQQGRISHTDI VIGCLVDLATGLMTFTANGKESNTFFQVEP TKLFPAVFVLFTHQNVIGFELGKKNIMPLS, AMFQSERKINPAQCPPRIEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDLELSERLDLQRHSHTLKL YRAVCALGNNRVAHALCSHVDQAQLLHAL DABILPGPLRAGYYDLLISHILESACRSRRSM SEVIPLTPETRATILFPGRSTENGHPRIGGI GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGGRAFBGG GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGGRAFBGGE EKEDEEETAQEKEDEEEEEBEABEGKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAPAER YVDKLQANQRSRYGLLIKK FSMTAAETARTREFFSPQEQNMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLHALGIQUL GEEEPPEEETTLOSKRUSELEEBLDLGRRSRSCH CFEEDEPLEFTTLOSKRUSHLEKVRLVKKKKE PEEERSAEESKPRSLQELVSHMVVRWAQED VQSPELVRAMFSLLHRQVDGLGHLAHGIQUL GEEEPPEEETTLOSKRUSHLEKVRLVKKKKE PEEERSAEESKPRSLQELVSHMVVRWAQED VQSPELVRAMFSLLHRQVDGLGLHALDF,CD TDEEDCPLPEEIRGDLLDFHQDLLHALGGRUD GEEEPPEETTLOSKRUSHLEKVRLVKKKKE PEEERSAEESKPRSLQELVSHMVVRWAQED VQSPELVRAMFSLLHRQVDGLGLHALDF,CD TDEEDCPLPEETRICHTLRYNGLLEKVRLVKKKE PEEERSAEESKPRSLQELVSHMVVRWAQED VQSPELVRAMFSLLHRQVDGLGELLRALPR, YTISPSSVEDTIMSLLEKVRLVKKKKE							ODDETO LOCADE EMBOSEED PROCESSOR TOTAL
LPPGYAPCHEAVLPREALHLEPIEVEYREGE RGPHLVGPSRCLSHTDPCPVDTVQVIVLPPI LERIREKLAENIHELWALTRIEQGWTYGPVR DNKRLHPCLVDFHSLPEPERRYNLOMSGETI KTLALAGCHYGMADEKAEDNLKKTKLPKT MMSNGYKPAPLDLSHVRLTPAQTTLVDRIA NGENVWAADRVGQGWSVAVQDIPARRNF LVPYRLLDEATKRSNRDSLCQAVRTLLGYG NIEPPDQEPSQVENQSRCDRVRIFAEKSYTY QSGRWYFEFEAVTTGEMRVGWARPELRPD ELGADELAYVFNOHRGQRWHLGSEPFGRPV QPGDVVGCMDLTENTHLNGEVLMSDSG ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQI VSSLRFFALCGLQEGFEPFANMQRPVTTWFF KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCL LTHRTWGSQNSLVEMLFLRLSLYQFHOHPI CTAGATPLAPPGLQPPAEDEARAAEPPDPDY NLRRSAGGWSEAENGKEGTAKEGAPGGTPF AGGEAQPAAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEILNTTTYYYSVRVFAGQEPSCWAGWVY PPYHQHDMSPDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQGRSHTDI VIGGLVDLATGLMTFLANKESNTFFQVEPP TKLFFAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMGNLMPVSWSI MPNHFLQVETRRAGERLGWAVQCCQPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALGSHDQAGLLHALL DAHLPGPLRAGYYDLLISHLESACSRRSM SEYIVPLTPETRAITLFPPGRSTENGHRPHGHGL GVGVTTSLRPPHHFSPFCVVALPAGAGAEA ARLSPAIPLEALRDKALRMLGGGDE DVKQILKMIEPEVTTEEEEEEDEEEGEEEGE EKEEDEET AQEKEDEKEEEGEEEGEEED EKEEDEET TAQEKEEDEEAEEGEEEGEE LEEGLLQMKLPSSVKLQANQRSRYGLLKK SMTAAETARTREFERSPQDINMLLQFKD TDEEDCPLPEEIRQDLLDHQDLLAHGQIOLI GEEEPPEEETTLOSRLMLLEKVALVKKEE PEEERSAEESKPRSLQELVSHMVVRWAQED VQSPELVRAMFSLLHRQVDGLELLRALPR, YTISPSSVEDTIMSLLELGLQIRSLLIVQMGPQ					1		
RGPHLVGPSRCLSHTDFVPCPVDTVQIVLPPI LERIREKLAENIHEL WALTRIEQGWTYGPVR DNKRLHPCLVDFHSLPFERNYNLQMSGETI KTLLALGCHVGMADEKAEDNLKKTKLPKTT MMSNGYKPAPLDLSTLYDRIAM NGHNVWARDRVGQGWSYSAVQDDPARRNY LVPYRLLDEATKRSNRDSLCQAVRTLLGYG NIEPPDQEPSQVENGSRCDRVRIPRAEKSYTY QSGRWYFEFEAVTTGEMRVGWARFELRPD ELGADELAYVPNGHGQWHLGSEPFGRPY QPGDVVGCMDLTENTIIFTLNGEVLMSDSG ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQI VSSLRFFALCGLGEGFEPFANMGRPVTTWFF KGLPQFEPVPLEHPHYEVSKYDGTVDTPPCL LTHRIWGSONSLVEMLFLRLSLPVQFHQHFI CTAGATPLAPPGLQPPAEDEARAAEPDPDYY NILRRSAGGWSEAENGKETAKEGAPGFTPY AGGEAQPARAENEKDATTEKNKKRGFLFLA KKVAMMTQPPATPTLPRLPHDVPADNRDI PEILNTTTYTYSVRVFAGQEPSCVWAGWY PDYHQHDMSFDLSKVRVVTVTMGDEQGHY HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMFTANGKESNTFPQVEPP TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPYSWSI MPNIFLLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDLELSERLDLQRFHSHTLRL YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGFLRAGYYDLLISHLESACRSRRSMM SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDLISHLESACRSRRSMS SEIVPLTPETRAITGAVTDCARGRYGLILKL SERESERSRRSLQALRSKALGRAGGEGEE LEEGLLQMKLPSSVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANGRRYGLILKL STATAATARRTREFSSPPQEQNMLLQFKD TDEEDCPLPEERIPQDLLDFHQDLAHCGIQLI GEEEPFEETTTLGSRLMSLLEKVRLVKKKE PEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYGGGGELLRALPR, YTISSSSVEDIMSLLECLGGRISLLIVQMGPQ						,	!
LERIREKLAENHEL.WALTRIEGGWTYGPYR DNKRLHPCLVDFHSLPEPERNYNLOMSGETI KTLLALGCHVGMADEKAEDNLKKTKLPKT MMSNGYKPAPLDLSHYRLTPAQTTLVDRIA NGHNYWARDRVGGWSYSAVQDPARRNE LVPYRLLDEATKRSNRDSLCQAVRTLLGYG NIEPPDQEPSQVENQSRCDRVGRFRAEKSYTY QSGRWYFEFEAVTTGEMRVGWARPELRPD ELGADELAYVFNGHRGQRWHLGSEPFGRPW QPGDVVGCMDLTEINTEITLIGEVLMSDSG ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQI VSSLRFFALCGLQEGFEFFANMQRPVTTWFF KGLPQFEPVPLEHPHYEVSRVDGTVOTPPCL LTHRTWGSONSLVEMI.FLRLSLPVQFHOHEI CTAGATPLAPPGLQPPAEDEARAAEPPDPY NIRRSAGGWSEAENGKEGTAKEGAPGGTPY AGGEAQPARAENEKDATTEKNKKRGFLFFA KKVAMMTQPPATPLRPHDVVPADNRDI PEILINTTTYYYSVRVFAGQEPSCVWAGWVT PDYHQHDMSFDLSKVRVVTYTMGDEQGNV HSSLKCSNCYMVWGGPFVSFQQQGRISHTD VIGCLVDLATGLMFTANGKESNTFFQVEPP TKLFPAAVFVLPTHOLYPELGKGKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNIHLQVETRRAGERLGWAVQCCQEPLTM ALHIPENRCMDLESENDLORFHSHTLRI YRAVCALGNNRVAHALCSHVDQAGLLHAL DAHLPGPLRAGYYDLLISHLESACRSRRSM SETVPLTPETRAITLFPFGRSTENGHPRHGLI GVGVTTSLRPPHIFSPPCFVAALPAAGAAEA ARLSPAPLLEALREKARM.GGAVROGGOPH CONTROL OF THE CO					·		LPPPGYAPCHEAVLPRERLHLEPIKEYRREGP
DNKRLHPCLVDFHSLPEPERNYNLQMSGED KTLALGCHVGMADEKAEDNLKKTKLPKT MMSNGYKPAPLDLSHVRLTPAQTTLVDRLA NGHNVWARDRUSGVBYSAVQDIPARRIN LVPYRLLDEATKRSNDSLCQAVRTILGYG NIEPPDQEPSQVENQSRCDRVRIPRAEKSYTY QSGRWYFEFEAVTTGEMRVGWARPELRPD ELGADELAYVPNGHRQRWHLGSEPFGRPV QPGDVVGCMIDLTENTIIETINGEVLMSDSG ETAFREIGIGGFLPVCSLGPGQVGHLNLGQI VSSLRFFALCIQ QEGFEPFADMQRPVTTWFP KGLPQFEVPLEHPHYEVCSLGPGQVGHLNLGQI VSSLRFFALCIQ QEGFEPFADMQRPVTTWFP KGLPQFPPPLEHPHYEVDTGTVDTPPCL LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFI CTAGATPLAPPGLQPPAEDEARAAEPDPDYY NLRSSAGGWSEAGMEGGTAKEGAPGGTP AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVPADNRDI PEILNTTTYYYSVENGAGEGSTAKEGAPGGTP HSSLKCSNCYMVWGGDFVSPGQQORISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEP TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPYSWSI MPNIFLQVETRAGERLGWAVQCQEPLTM ALHPEENRCMDLESERLDJ QRFHSHTLRL YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGFLRAGYYDLLISHLESACRSRRSMSM SEYVPLTPETRAITESRLDJ AQRHSHTLRL TRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGFLRAGYYDLLISHLESACRSRRSMSM SEYVPLTPETRAITEPRGSTENGHPRIGLI GVGVTTSLRPPHIFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDFVGASVEPQFYPVLKLVSTILLVMGIFGGDE DVKQULKMIEPSVFTEEEEEEDEEEEGEEEDE EKEEDEEETAGEKEDEEEGEEEDEE EKEEDEETAGEKEDEEEGEEEDEE EKEEDEETAGEKEDEEEGEEEDEE EKEEDEETAGERKDEKEEFEAAGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIKL FSHTAAETARTREFSPPQEQNMLLQFKD TDEEDCPLPEEIRQULLDFHQDLAHCIGIQLI GEEEFFEETTIGSRLMSLLEKVRLVKKKE PEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYGGIGELLRALPR. YTISPSSVEDIMSLLECLGQRRSLLIVKJKKE							RGPHLVGPSRCLSHTDFVPCPVDTVQIVLPPH
KTLLALGCHVGMADEKAEDNLKKTKLPKT MMSNGYKPAPLDLSHVRLTPAQTTLVDRLA NGHNVWARDRVGQWSYSAVQDIPARRNP LVPYRLLDEATKRSNRDSLQAVWITLLGYG NIEPPDQEPSQVENQSRCDRVBFRAEKSYTY QSGRWYFEFEAVTTGEMRVGWARPELRPDV ELGADELAYVFNOHRGGGRWHLGSEPFGRPV GPGDVVGCMDLTENTHTINTGSVLMSDSG ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQI VSSLRFAICGLQEGFEPFARMQRPVTTWFS KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCL LTHRTWGSQNSLVEMLFLR.SLPVQFHQHFI CTAGATPLAPPGLPPAEDEARAAFPDPDYY NLRRSAGGWSEAENGKEGTAKEGAPGGTPC AGGEAQPARAENEKDATTTEKNKKRGFLFFA KKVAMMTQPATPTLPRLPHDVYADNRDI PEIILNTTTYYSVRVFAGQEPSCVWAGWV PDYYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDPVSPQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEPN TKLFPAVFVLPTHQNVIQFELGKKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHPEENRCMDILELSERLDLQRFHSITLIKL YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYDLLISHLESACRSRRSMI SEYIVPLTPETRAITLFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCTVAALPAAGAAEA ARLSPAPLEALBRKALRMLGEAVRDGGQH RDPVGASVEFQFYPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEDLEEEEGEEEEGE EKEEDEETAGEKEDEKKEEERAAFGKEEC LEEGLLQMKLPESVKLQMCHLLEYFCOQEL HRVESLAAPAERAYVDKLQANQKSRYGLIK FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDTHQDLLAHCGIQLI GEEEPEEPETILGSRLMSLLEKURJUKKKE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQVDGLIGELLRALPR YTISPSSVEDIMSSLLECLGQIRSLLIVVQMGPD VQSPELVRAMFSLLHRQVDGLIGLLALPR YTISPSSVEDIMSSLLECLGQIRSLLIVVQMGPD							LERIREKLAENIHELWALTRIEQGWTYGPVRD
MMSNGYKPAPL.DLSHVRI.TPAQTTI.VDRI.A NGHNVWARDRVGQGWSYSAVQDIPARRINF LVPYRLLDEATKRSNRDSI.CQAVRTILIGYG NIEPPDQEPSQVENQSRCDRVBIFRAEKSYTV QSGRWYFEFEAVTIGEMRVGWARPELRPDV ELGADELAYVFNOHRGGRWHLGSEPFGRPV QPGDVVGCMDLTENTIITITINGEVLMSDSG ETAFREIEIGDGFLPVCSLGPGQVGH.NLGQ) VSSLRFFAICGLQEGFEPFAINMQRPVTTWFS KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCL LTHRTWGSONSL.VEMLFLRI.SLPVQFHOHFI CTAGATPLAPPGLQPPAEDEARAACEPDPDYY NLRRSAGGWSEAENGKEGTAKEGAFAGEPPDYY NLRRSAGGWSEAENGKEGTAKEGAFAGEPPDYY NLRRSAGGWSEAENGKEGTAKEGAFAGEPPDYY MIRRSAGGWSEAENGKEGTAKEGAFAGEPPDYY PSILNTTTYYYSVRVFAGQEPSCVWAGWVT PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGINFTANGKESNTFFQVEPP TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPLEMQMLMPVSWSI MPNHFLQVETRAGRELGWAVQCQPLTMA ALHPEENRCMDLELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISHLESACRSRRSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAPLEALARDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGGDE DVKQILKMIPPEVTTEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEKEEEBAAEGEKEE LEEGLLQMKLPESVKLQWCHLLEYFCDQEL HRVESLAAFAERFVDKLQANQRSRYGLIKK FSMTAAETARRTREFRSPPQEQNMLLQFKD TDEEDCPLPEEIRQDLLDHTQDLLAHCGIQLI GEEEEPEETTILGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQED VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQUSLLIVVQMGPQ	Į į						DNKRLHPCLVDFHSLPEPERNYNLQMSGETL
NGHNYWARDRYGGGWSYSAVQDIPARRIFI LVPYRLLDEATKRSNRDSLCQAVRTILLGYG NIEPPPOGEPSQVENQSRCDRYWIPRAEKSYTY QSGRWYFEFEAVTTGEMFVGWARPELRPD ELGADELAYVFNGHRGGRWHLGSEPFGRPY QFGDVVGCMIDLTENTIIFTLNGEVLMSDSG ETAFREIEIGDGFLPYCSLGPGQVGHLNLGQI VSSLRFALGGLGGFEPFANNQRPYTTWFF KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCL LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFI CTAGATPLAPPGLQPPAEDEARAAEPDPDYY NILRSAGGWSEAENGKEGTAKEGAPGGTPC AGGEAQPARAENEADATTENKKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEILINTTTYYYSVRVFAGGEPSCVWAGWV PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEPN TKLFPAVFVLFTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRAGERLGWAVQCOPPLTM ALHIPEENRCMDLELSERLDLQRFHSHTLRL YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLFGPLRAGYDLISHLESACRSRRSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCTYAALPAGAALA ARLSPAPLEALARMKARMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGGB DVKQILKMIEPEVFTEEEEEDEEEEGEEEDE EKEEDEETAGEKEDEKKEEERAAEGEKEEC LEEGLLQMKLPESVKLQMCHLLEYFCOQEL HRVESLAAFAERRYVDKLQANGRSRYGLLIK, FSMTAAETARRTEFFSPPQEQNMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEPETITGSRLMSLLEKURJVKKEE PEEERSAEESKPRSLQELVSHMYVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQUSLLIVQMGPQ	1 '						KTLLALGCHVGMADEKAEDNLKKTKLPKTY
NGHNYWARDRYGGGWSYSAVQDIPARRIFI LVPYRLLDEATKRSNRDSLCQAVRTILLGYG NIEPPPOGEPSQVENQSRCDRYWIPRAEKSYTY QSGRWYFEFEAVTTGEMFVGWARPELRPD ELGADELAYVFNGHRGGRWHLGSEPFGRPY QFGDVVGCMIDLTENTIIFTLNGEVLMSDSG ETAFREIEIGDGFLPYCSLGPGQVGHLNLGQI VSSLRFALGGLGGFEPFANNQRPYTTWFF KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCL LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFI CTAGATPLAPPGLQPPAEDEARAAEPDPDYY NILRSAGGWSEAENGKEGTAKEGAPGGTPC AGGEAQPARAENEADATTENKKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEILINTTTYYYSVRVFAGGEPSCVWAGWV PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEPN TKLFPAVFVLFTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRAGERLGWAVQCOPPLTM ALHIPEENRCMDLELSERLDLQRFHSHTLRL YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLFGPLRAGYDLISHLESACRSRRSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCTYAALPAGAALA ARLSPAPLEALARMKARMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGGB DVKQILKMIEPEVFTEEEEEDEEEEGEEEDE EKEEDEETAGEKEDEKKEEERAAEGEKEEC LEEGLLQMKLPESVKLQMCHLLEYFCOQEL HRVESLAAFAERRYVDKLQANGRSRYGLLIK, FSMTAAETARRTEFFSPPQEQNMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEPETITGSRLMSLLEKURJVKKEE PEEERSAEESKPRSLQELVSHMYVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQUSLLIVQMGPQ							MMSNGYKPAPLDLSHVRLTPAOTTLVDRLAE
LVPYRLLDEATKRSNRDSLCQAVRTILIGYG NIEPPDQEPSQVENQSRCDRVRIFRAEKSYTV QSGRWYFEFEAVTTOEMRVGWARPELRPD ELGADELAYVFNGHRGQRWHLGSEPFGRPV QPGDVVGCMIDLTENTIIFTLNGEVLMSDSG ETAFREIEIGDFLPVCSLGFQVGHLNLGQI VSSLRFFAICGLQEGFEPFAINMQRPVTTWFF KGLPOFEPVPLEHPHYEVSRVDGTVDTPPCL LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFI CTAGATPLAPPGLQPPADEARAAEPDPDYY NLRRSAGGWSEAENGKEGTAKEGAPGGTPY AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEIILNTITTYYSVRVFAGQEPSCVWAGWV PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTPTANKESNTFFQVEPY TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRAGERLGWAVQCQEPLTM ALHIPEENRCMDLELSERLDLQFFHSHTLRI YRAVCALGNINVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISHLESACRSRRSM SEYIVPLTPETRATILEPPGRSTENGHPHGLI GVGVTTSLRPPHHFSPPGFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFPVLKLVSTLLVMGFFGDE EKEEDEETAQEKEDEEKEEEEAEEGEEDEE EKEEDEETAQEKEDEEKEEEEAAEGEKEE LEEGLLQMKLPESVKLQMCHLLEYTCDQEL HRVESLAAFAAERYVDKLQANQRSRYGLLK FSMTAAETARRTREFRSPPQEQNMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEERSAEESKPRSLQELUSHMVVRWAQED VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQIRSLLIVQMGPGP							
NIEPPDQEPSQVENQSRCDRVRIFRAEKSYTT QSGRWYFEFEAVTTGEMRVGWARPELRPDP ELGADELAYVPNGHRGQRWHLGSEPFGRPV OPGDVVGCMIDLTENTIIFTLNGEVLMSDSG ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQI VSSLRFFAICGLQEGFEPFAINMQRPVTTWFS KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCL LTHRTWGSQNSLVEMIFLRLSLPVQFHQHFFI CTAGATPLAPPGLQPPAEDEARAAEPDPDYS NLRRSAGGWSEAENGREGTAREGAPGGTPP AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEIILINTTTYYYSVRVFAGQEPSCVWAGWVY PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEPP TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKINPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGEBLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRL YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGFLRAGYYDLLISHLESACRSRSSMI SEYIVPLTPETRATILFPPGRSTENGHPRHGLI GVGVTTSLRPHHFSPPCFVAALRAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEPGFVPVLKLVSTLLVMGIFGDE EKEEDEEFTAQEKEDEEEGEEEEDEEGEGEEDE EKEEDEEFTAQEKEDEEKEEEBAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HIRVESLAAFAERYVDKLQANQRSRYGLLIK FSMTAAETARRTREFRSPPQEQINMLOFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEEETTLGSRLMSLLEKVRLVKKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQED VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQRSLLIVQMGPQ]						
QSGRWYFEFEANTTGEMRVGWARPELRPD ELGADELAYVFNGHEGQRWHLGSEPFGRPV QPGDVVGGMDLTENTIIFTLNGEVLMSDSG ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQI VSSLRFFAICGLQEGFEPFAINMQRPVTTWFF KGLPOFEVPYLEHPHYEVSRVDGTVDTPPCL LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFI CTAGATPLAPPGLQPPADEARAAEPDPDYT NILRSAGGWSEAENGKEGTAKEGAPGGTPY AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEIILINTTTYYSVRVFAGGEPSCVWAGWVY PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEPN TKLIFPAVFVLPHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPYSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDLLELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLJHAL DAHLPGPLRAGYYDLLISHLESACRSRSM SEYIVPLTPETRATILFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVDGQGH RDPPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQLKMIEPEVFTEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEEEAAEGEKEEGL LEEGLLQMKLPESVKLQMCHLLEYFCQOEL HRVESLAAFAERYVDKLQANQRSRYGLIK FSMTAAETARRTREFRSPPQEQINMLQFKD TDEEDCPL-PEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEETTLGSRLMSLLEKVRLVKKKEE PEERSAEESKFRSLQELVSHMVVRWAQED VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDITMSLLECLGQRSLLIVQMGPQ	!						
ELGADELAYVFNGHRGQRWHLGSEPFGRPV QPGDVVGCMIDLTENTIIFTLNGEVLMSDSG ETAFREIEIGGFLPVCSLGPGQVGHLNLGQI VSSLRFALCGLQEGFEPFAINMQRPVTTWFE KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCL LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFI CTAGATPLAPPGLQPPAEDEARAAEPDPDYF NLRRSAGGWSEAERGKEGTAKEGAFGGTFV AGGEAQPARAENEKDATTKENKKRGFLFAA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PPILINTITTYYSVRYAGQEPSCVWAGWVY PPVHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHID VIGCLVDLATGLMTFTANGKESNTFFQVEFP TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGFLRAGYYDLLISHLESACRSRRSM SEVIVPLTFETRATILFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGQH RDPYGASVEFQFVPVLKLVSTLLVMGIFGDB DVKQILKMIEPEVFTEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEEEAAEGGKEEC LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK FSMTAAETARRTREFRSPPQEQNMLLOFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPFEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDJ VQSBUVSHMVVRWAQEDJ VTSPSSVEDIMSLLECLGQIRSLLIVQMGPQ	-			i			
QPGDVVGCMIDLTENTIITLNGEVLMSDSG ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQI VSSLRFFALCGLQEGFEPFANMQRPVTTWFS KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCL LTHRTWGSQNSLVEMLFLRLSLPVQFHQHHFI CTAGATPLAPPGLQPPAEDBEARAAEPDPDYS NLRRSAGGWSEAENGKEGTAKEGAPGGTPC AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEILINTTTYYSVRVFAGQEPSCVWAGWVY PPVHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQRISHTD VIGGLVDLATGLMTFTANGKESNTFFQVEPP TKLFPAVFVLPTHQNVQFELGKQKNIMPLS, AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNIFLQVFTRRAGERLGWAVQCQFPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLAGYYDLLISHILESACRSRSM, SEYIVPLTPETRAITLFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCTVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEQFVPVLKLVSTILLVMGIFGDE DVKQILKMIEPEVFTEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEEELAAEGGKEEC LEEGGLQMKLPESVKLQMCHILEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK FSMTAAETARRTREFRSPPQEQINMILQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDU VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTSPSSVEDIMSLLECLGQIRSLLIVQMGFQ							
ETAPREIEIGDGFLPVCSLGPGQVGHLNLGQI VSSLRFFAICGLQGEFEPFAINMQRPVTTWFS KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCL LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFI CTAGATPLAPPGLQPPAEDEARAAEPDPDYY NLRRSAGGWSEAENGKEGTAKEGAPGGTPC AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEIILNTTTYYSRVFAGQEPSCVWAGWV PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEPN TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHISHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISHLESACRSRRSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLE GVGVTTSLRPPHHFSPPCTVAALPAAGAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQLKMMEPEVFTEEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEEEEAAEGEKEEC LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLK FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLDHCGIGLL FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRGDLLDFHQDLLAHCGIGUL GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEERSAEESKPRSLQELVSHMVVRWAQED VQSPELVAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQIRSLLIVQMGPG							f
VSSLRFFAICGLOEGFEPFAINMQRPVTTWES KGLPQFEPVPLEHPHYEVSVDGTVDTPPCL LTHRTWGSQNSLVEMLFLR.SLPVQPHQHFI CTAGATPLAPPGLQPPAEDEARAAEPDPDYF NLRRSAGGWSEAENGKEGTAKEGAPGGTPC AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEIILNTTTYYSVRVFAGQEPSCVWAGWV PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEPN TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNINRVAHALCSHVDQAQLLHAL DAHLPGPIRAGYYDLLISHLESACRSRRSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEGEEEED EKEEDEETAQEKEDEEKEEEEAAEGEKEEC LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK FSMTAAETARRTEFRSPPGQONMLJQFKD TDEEDCPLPEEIRQDLLDFHQDILLAHCGIGUL FFSMTAAETARRTEFRSPPGQONMLJQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIGUL GEEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQED VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPG		,					1?
KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCL LTHRTWGSQNSLVEMIFLRLSLPVQFHQHFFI CTAGATPLAPPGLQPPAEDEARAAAEPDPDYS NLRRSAGGWSEAENGKEGTAKEGAPGGTPC AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEILLNTTTYYYSVRVFAGQEPSCVWAGWV' PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEPP TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MFNHFLQVETTRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGFLRAGYYDLLISHLESACRSRSSM SEYIVPLTPETRAITLFPPGRSTENGHPPHGLI GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEDEEEGEEEDE EKEDDEETAQEKEDEEKEEEEAAEGEKEEC LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPPEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDJ VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDIMSLLECLGQIRSLLIVQMGPQ	1						
LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFI CTAGATPLAPPGLQPFAAEDARAAAEPDPDYF NLRRSAGGWSEAENGKEGTAKEGAPGGTPC AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEILINTTTYYYSVRVFAGQEPSCVWAGWVT PPYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEPF TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRL YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISHLESACRSRRSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQDLKMIEPEVFTEEEEEDEEEGEEDEDE EKEDDEETAQEKEDEEKEEEEAAEGEKEEC LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPPEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDITMSLLECLGQIRSLLIVQMGPQ							VSSLRFFAICGLQEGFEPFAINMQRPVTTWFS
CTAGATPLAPPGLQPPAEDEARAAEPDPDYT NILRRSAGGWSEAENGKEGTAKEGAPGGTTPC AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEIILNTTTYYYSVRVFAGQEPSCVWAGWVT PDYHQHDMSFDLSKVRVVTVTMGDEQGNVV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMITFTANGKESNTFFQVEPP TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS, AMFQSERKNAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISHHLESACRSRRSMI SEYIVPLTPETRATITLFPPGRSTENGHPRHGIL GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVVVLKLVSTLLVMGIFGDB DVKQILKMIEPEVFTEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHILLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIK, FSMTAAETARRTREFRSPPOQQINMILLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAESKPRSLQELVSHMVVRWAQEDJ VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ	l I						KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCLR
CTAGATPLAPPGLQPPAEDEARAAEPDPDYT NILRRSAGGWSEAENGKEGTAKEGAPGGTTPC AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEIILNTTTYYYSVRVFAGQEPSCVWAGWVT PDYHQHDMSFDLSKVRVVTVTMGDEQGNVV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMITFTANGKESNTFFQVEPP TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS, AMFQSERKNAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISHHLESACRSRRSMI SEYIVPLTPETRATITLFPPGRSTENGHPRHGIL GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVVVLKLVSTLLVMGIFGDB DVKQILKMIEPEVFTEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHILLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIK, FSMTAAETARRTREFRSPPOQQINMILLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAESKPRSLQELVSHMVVRWAQEDJ VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ							LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFR
NLRRSAGGWSEAENGKEGTAKEGAPGGTPC AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEILINTTTYYYSVRVFAGQEPSCVWAGWV' PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEFP TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS, AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISHILESACRSRRSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLE GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEFEAAEGEKEEC LEEGLLQMKLPESVKLQMCHILLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIK FSMTAAETARRTREFRSPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEFPEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQED VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ							
AGGEAQPARAENEKDATTEKNKKRGFLFKA KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEILINTTTYYYSVRVFAGQEPSCVWAGWY PPYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVVDLATGLMTFTANGKESNTFFQVEPN TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS, AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISHLESACRSRSRMSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEDEEEGEEEDEE EKEEDEETAQEKEDEEKEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK, FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ							
KKVAMMTQPPATPTLPRLPHDVVPADNRDI PEIILNTTTYYYSVRVFAGQEPSCVWAGWVI PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEPN TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS, AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISIHLESACRSRRSMI SEYIVPLTPETRAITLFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQELL HRVESLAAFAERYVDKLQANQRSRYGLIK, FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDJ VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ	ł I						
PEIILNTTTYYYSVRVFAGQEPSCVWAGWVY PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYWWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEPN TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS, AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLHAL DAHLPGPLRAGYYDLLISIHLESACRSRRSMI SEYIVPLTPETRAITLFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVVVLKI.VSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEDEEEGEEEDE EKEEDEEETAQEKEDEEKEEEAAEGEKEEC LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ							
PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEFN TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS, AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISHLESACRSRRSMI SEYIVPLTPETRAITLFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEDEEEEGEEDED EKEEDEETAQEKEDEEKEEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIK, FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPR, YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ							
HSSLKCSNCYMVWGGDFVSPGQQGRISHTD VIGCLVDLATGLMTFTANGKESNTFFQVEPN TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS, AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISIHLESACRSRRSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLE GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEGEEEDE EKEEDEETAQEKEDEEKEEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIK, FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDILAHCGIQLI GEEEPEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPR, YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ	į						l
VIGCLVDLATGLMTFTANGKESNTFFQVEPN TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISIHLESACRSRRSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLE GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDB DVKQILKMIEPEVFTEEEEEEDEEEEGEEDE EKEEDEEETAQEKEDEEKEEEEAAEGEKEE LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEIRQDLLDFHQDLLAHCGIQLI GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ	1		1				
TKLFPAVFVLPTHQNVIQFELGKQKNIMPLS. AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISIHLESACRSRRSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLE GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEGEEEDE EKEEDEEFTAQEKEDEEKEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA						*	
AMFQSERKNPAPQCPPRLEMQMLMPVSWSI MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISIHLESACRSRRSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLE GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEGEEEDE EKEEDEETAQEKEDEEKEEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIK FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA							l •
MPNHFLQVETRRAGERLGWAVQCQEPLTM ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISIHLESACRSRRSM SEYIVPLTPETRAITLFPPGRSTENGHPRHGLI GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEGEEEDE EKEEDEETAQEKEDEEKEEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIKL FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA							
ALHIPEENRCMDILELSERLDLQRFHSHTLRI YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISIHLESACRSRRSMI SEYIVPLTPETRAITLFPPGRSTENGHPRHGLE GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEGEEEDE EKEEDEETAQEKEDEEKEEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIKL FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ							
YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISIHLESACRSRRSMI SEYIVPLTPETRAITLFPPGRSTENGHPRHGLE GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEGEEEDE EKEEDEETAQEKEDEEKEEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIKL FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ	1						
YRAVCALGNNRVAHALCSHVDQAQLLHAL DAHLPGPLRAGYYDLLISIHLESACRSRRSMI SEYIVPLTPETRAITLFPPGRSTENGHPRHGLE GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEGEEEDE EKEEDEETAQEKEDEEKEEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIKL FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ					·		ALHIPEENRCMDILELSERLDLQRFHSHTLRL
DAHLPGPLRAGYYDLLISIHLESACRSRRSMI SEYIVPLTPETRAITLFPPGRSTENGHPRHGLE GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEGEEEDE EKEEDEETAQEKEDEEKEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIKA FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ	i '						YRAVCALGNNRVAHALCSHVDQAQLLHALE
SEYIVPLTPETRAITLFPPGRSTENGHPRHGLE GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEEEAAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIKL FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ	1						DAHLPGPLRAGYYDLLISIHLESACRSRRSML
GVGVTTSLRPPHHFSPPCFVAALPAAGAAEA ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHILLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK, FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ		,			1		
ARLSPAIPLEALRDKALRMLGEAVRDGGQH RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHILLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK, FSMTAAETARRTREFRSPPQEQINMILQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ							
RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE DVKQILKMIEPEVFTEEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHILLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK, FSMTAAETARRTREFRSPPQEQINMILQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ							
DVKQILKMIEPEVFTEEEEEEDEEEEGEEEDE EKEEDEETAQEKEDEEKEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLIIK, FSMTAAETARRTREFRSPPQEQINMILLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDJ VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ							
EKEEDEETAQEKEDEEKEEEAAEGEKEEG LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK, FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ	1						
LEEGLLQMKLPESVKLQMCHLLEYFCDQEL HRVESLAAFAERYVDKLQANQRSRYGLLIK, FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ	1				1		
HRVESLAAFAERYVDKLQANQRSRYGLLIK. FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ							
FSMTAAETARRTREFRSPPQEQINMLLQFKD TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ						İ	
TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ			1		}		HRVESLAAFAERYVDKLQANQRSRYGLLIKA
TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLI GEEEEPEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ							FSMTAAETARRTREFRSPPQEQINMLLQFKDG
GEEEEPEETTLGSRLMSLLEKVRLVKKKEE PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ						,	TDEEDCPLPEEIRODLLDFHODLLAHCGIOLD
PEEERSAEESKPRSLQELVSHMVVRWAQEDI VQSPELVRAMFSLLHRQYDGLGELLRALPR YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ					1		GEEEPEETTLGSRLMSLLEKVRI VKKKEEK
VQSPELVRAMFSLLHRQYDGLGELLRALPRA YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ					ľ		
YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQ							
							V V O T EL V KAN IF OLL TIK V I DOLU ELL KALPKA
TO THE REPORT OF THE PROPERTY							
ENLMIQSIONIMANA VET QHENLMRALGMH	1				}		ENLMIQSIGNIMNNKVFYQHPNLMRALGMHE
TVMEVMVNVLGGGESKEIRFPKMVTSCCRF	L				L		TVMEVMVNVLGGGESKEIRFPKMVTSCCRFL

peptide /=possible nucleotide deletion, \=p nucleotide insertion CYFCRISRQNQRSMFDHLSYLI QGSTPLDVAAASVIDNNELAL, VSYLAGCGLQSCPMLVAKGYI ERYLDFLRFAVFVNGESVEEN, KPECFGPALRGEGGSGLLAAIE DGPGIRRDRRREHFGEEPPEEN	id, Histidine, le, Proline, ine, tophan, ocodon, ossible
QGSTPLDVAAASVIDNNELAL. VSYLAGCGLQSCPMLVAKGYI ERYLDFLRFAVFVNGESVEEN. KPECFGPALRGEGGSGLLAAIE DGPGIRRDRRREHFGEEPPEEN	
VSYLAGCGLQSCPMLVAKGYI ERYLDFLRFAVFVNGESVEEN KPECFGPALRGEGGSGLLAAIE DGPGIRRDRREHFGEEPPEEN	
KPECFGPALRGEGGSGLLAAIE DGPGIRRDRREHFGEEPPEEN	
DGPGIRRDRRREHFGEEPPEEN	
FYAALIDLLGRCAPEMHLIQA(
LRSLVPLEDLVGIISLPLQIPTLC MSASFVPDHKASMVLFLDRV	
VLDVGFLPDMRAAASLDTATE	FSTTEMALAV
NRYLCLAVLPLITKCAPLFAGT MLHTVYRLSRGRSLTKAQRD	
YIRPSMLQHLLRRLVFDVPILN	TEFAKMPLKLL
TNHYERCWKYYCLPTGWANF TRKLFWGIFDSLAHKKYDPEL	
AIAGALPPDYVDASYSSKAEK	KATVDAEGNF
DPRPVETLNVIIPEKLDSFINKF DKIONNWSYGENIDEELKTHP	
KDKÈIYRWPIKESLKAMIAWE	WTIEKAREGE
EEKTEKKKTAKISQSAQTYDPI SAVTLSRELQAMAEQLAENYF	•
ELEAKGGGTHPLLVPYDTLTA	
QELLKFLQMNGYAVTRGLKD	
FAFGFLQQLLRWMDISQEFIAI EKSPHEQEIKFFAKILLPLINQY	
TPAKVLGSGGHASNKEKEMIT	SLFCKLAALV
RHRVSLFGTDAPAVVNCLHIL, KSGPEIVKAGLRSFFESASEDII	
KVSQARTQVKGVGQNLTYTT	
QHIAQHQFGDDVILDDVQVSC TTKNTYVEKLRPALGECLARL	
PQLNEYNACSVYTTKSPRERA	ILGLPNSVEEM
CPDIPVLERLMADIGGLAESGA ITLPMLCSYLPRWWERGPEAP	
CTAVTSDHLNSLLGNILRIIVN	NLGIDEASWM
KRLAVFAQPIVSRARPELLQSI AGKVVSEEEQLALEAKAEAQI	
VLCRDLYALYPLLIRYVDNNR	AQWLTEPNPS
AEELFRMVGEIFIYWSKSHNFI NEINNMSFLTADNKSKMAKA	
RTKKKRRGDRYSVQTSLIVAT	LKKMLPIGLN
MCAPTDQDLITLAKTRYALKI NNLHLQGKVEGSPSLRWQMA	
DADDPEKIVRRVQEVSAVLYY	/LDQTEHPYKS
KKAVWHKLLSKQRRRAVVAQ THRACNMFLESYKAAWILTEI	
LSKAGEQEEEEEVEEKKPDP	
ALTEKSKLDEDYLYMAYADI GENGEAEEEVEVSFEEKQMER	
HTRGAAEMVLQMISACKGET	
GISILNGGNAEVQQKMLDYLK	DKKEVGFFQS
IQALMQTCSVLDLNAFERQNK DGTVINRQNGEKVMADDEFT	QDLFRFLQLLC
EGHNNDFQNYLRTQTGNTTT	NIIICTVDYLL
RLQESISDFYWYYSGKDVIEEG SVAKQVFNSLTEYIQGPCTGN	
DAVVGFLHVFAHMMMKLAQ	DSSQIELLKEL
LDLQKDMVVMLLSLLEGNVV DMLVESSSNVEMILKFFDMFL	
QDYVTDPRGLISKKDFQKAMI	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide	-	/=possible nucleotide deletion, \=possible
		İ		sequence		nucleotide insertion
						QFLLSCSEADENEMINCEEFANRFQEPARDIG
			1			FNVAVLLTNLSEHVPHDPRLHNFLELAESILE
						YFRPYLGRIEIMGASRRIERIYFEISETNRAQW
		l		ŀ		EMPQVKESKRQFIFDVVNEGGEAEKMELFVS
	į		ļ			FCEDTIFEMQIAAQISEPEGEPETDEDEGAGA
	}					AEAGAEGAEGAAGLEGTAATAAAGATARV
			}			VAAAGRALRGLSYRSLRRRVRRLRRLTAREA
		1		1		ATAVAALLWAAVTRAGAAGAGAAAGALGL
	1	ĺ				LWGSLFGGGLVEGAKKVTVTELLAGMPDPT
						SDEVHGEQPAGPGGDADGEGASEGAGDAAE
						GAGDEEEAVHEAGPGGADGAVAVTDGGPFR
		1				PEGAGGLGDMGDTTPAEPPTPEGSPILKRKLG
						VDGVEEELPPEPEPEPEPELEPEKADAENGEK EEVPEPTPEPPKKQAPPSPPPKKEEAGGEFWG
			1	1		ELEVQRVKFLNYLSRNFYTLRFLALFLAFAIN
				<u> </u>		FILLFYKVSDSPPGEDDMEGSAAGDVSGAGS
]					GGSSGWGLGAGEEAEGDEDENMVYYFLEES
		Į				TGYMEPALRCLSLLHTLVAFLCIIGYNCLKVP
						LVIFKREKELARKLEFDGLYITEQPEDDDVKG
]		QWDRLVLNTPSFPSNYWDKFVKRKVLDKHG
		İ				DIYGRERIAELLGMDLATLEITAHNERKPNPP
			ĺ	•		PGLLTWLMSIDVKYQIWKFGVIFTDNSFLYLG
						WYMVMSLLGHYNNFFFAAHLLDIAMGVKTL
	Ì					RTILSSVTHNGKQLVMTVGLLAVVVYLYTVV
		1				AFNFFRKFYNKSEDEDEPDMKCDDMMTCYL
						FHMYVGVRAGGGIGDEIEDPAGDEYELYRVV
						FDITFFFFVIVILLAIIQGLIIDAFGELRDQQEQV
						KEDMETKCFICGIGSDYFDTTPHGFETHTLEE
						HNLANYMFFLMYLINKDETEHTGQESYVWK
						MYQERCWDFFPAGDCFRKQYEDQLS
501	1851	Α	3869	467	665	VIVAIYCQLIFDKGAKTIQ*PFQQIAL/CKRMK
	ļ	1				LGPCFTPCGKINSEWIRELSVRVKTIKHLEIGV
500	1050	<u> </u>	0000			N .
502	1852	Α	3888	1042	724	SGMQWRDLTPLQPLPPRFKQFSCLSLPGSWD
	l	1	ł	1	1	YRHAP\PLLTNF*FLVEMGFCYVGQAGRKLL
	i					ASSDQSALASQSAGITGISTAPGPPFFFLNFEA
503	1853	A	3891	1772	1102	GSCSVAQAGVQ
203	1023	^	1695	1773	1193	EVDSQSGVQ*QAPGSLQLQTPGLK/VSCLLSR
	[QDYRSSLPHLASCCYYYYYY/VFL*RRGLTTL VQGGLKLLPSSNPFASAP*TAGITGMSHCAGP
	1		1			HFNF*MFRKISCIRE*F*HTRIYDIPFLILFFKET
		1]		}	WVLLCYPGWPOIPGLKPSSCLRLLSSWDHRC
	1					APPCPASFFIFHVDRVSPPCPGLVSITFKMLLL
						L
504	1854	В	3896	279	70	MVSKSKSILMSYNHVELTFSDMKKMPEAFRR
	1007	~	5050	1	l '``	TQKHTTYLIPYQVIFWSTGKDAMRSFMMPFY
	l	l	1	1		QKEYYENQ*
505	1855	A	3899	2	1396	EPGVPTKKTWFDKPDFNRTNSPGFQKKVQFG
				-		NENTKLELRKVPPELNNISKLNEHFSRFGTLV
			}			NLOVAYNGDPEGALIQFATYEEAKKAISSTEA
						VLNNRFIKVYWHREGSTQQLQTTSPKVMQPL
	l	1	}	1		VQQPILPVVKQSVKERLGPVPSSTIEPAEAQS
	1		1			ASSDLPQVLST\LLA*QKQCIIQLL/WKAAQKT
	I .		1			LLVSTSAVDNNEAQKKKQEALKLQQDVRKR
	ļ	ł		ì		
						KOEILEKHIETOKMLISKLEKNKTMKSEDKAE
						KQEILEKHIETQKMLISKLEKNKTMKSEDKAE
				*		KQEILEKHIETQKMLISKLEKNKTMKSEDKAE IMKTLEVLTKNITKLKDEVKAASPGRCLPKSI KTKTQMQKELLDTELDLYKKMQAGEEVTEL
				*	-	KQEILEKHIETQKMLISKLEKNKTMKSEDKAE IMKTLEVLTKNITKLKDEVKAASPGRCLPKSI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NC: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion AFTESDREDLLPHFAQYGEIEDCQIDDSSLHA VITFKTRAEAEAAAVHGARFKGQDLKLAWN
506	1856	A	3911	1952	919	KPVTNISAVETEEVEPDEEEQREIIIA DAELSGTLSLVLTQCCKRIKDTVQKLASDHK DIHSSVSRVGKAIDKNFDSDISSVGIDGCWQA DSQRLLNEVMVEHFFRQGMLDVAEELCQES GLSVDPSQKEPFVELNRILEALKVRVLRPALE WAVSNREMLIAQNSSLEFKLHRLYFISLLMG GTTNQREALQYAKNFQPFALNHQKDIQVLM GSLVYLRQGIENSPYVHLLDANQWADICDIFT RDACALLGLSVESPLSVSFSAGCVALPALINIK AVIEQRQCTGVWNQKDELPIEVVDLG*KSAGY HSIFACPILRQQTTDNNPPMKLVCGHIISRDAL NKMFNGSKLKCPYCPMEQSPGDAKQIFF
507	1857	A	3936	439	18	SHPFSPAPGICPDAPPPLPRPSKGLGHPGTAGA PGSGARCHPPSTCSPSWASPG*GAKASPALPR SHGVTLLCKAQAHLCRGEDSKDASGSTSQA WEPG*GAWGMPRCQGPALGSCFCPPGTTVQ RPAKORDKRNRHLGR
508	1858	A	3944	120	412	WCPAGTLDFPGPQEMVLLEIEVMNQLNHRNL IQLYAAIETPHEIVLFME\YECPK*W*GLGGGT TRHGASRGGVCAHSIEGGELFERIVDEDYHLT EV
509	1859	A	3949	31	392	LTKTPSPREKGRGVLSVLLMMI*KCRVIFVKIP MVFFLQNFC/RIILNVA\WTGD*PNTL*KEQRG ITFSDSKS*YKATKIKTMWYCHKNRYID/ERN RIEIPEINPCICDKIIFRKLSMTTQ
510	1860	A	3954	1013	885	FSETRACCPRLEHSGRIEAHCSLNIPGSSDPPT SASSVAATTG
511	1861	A	3956	1	1054	PPAWAPRSPLIWAPTSGRHPCRAALPWSTSSV RWQPSEKQPPPPAHRGPADSLSTAAGAAELS AEGAGKSRGSGEQDWVNRPKTVRDTLLALH QHGHSGPFESKFKKEPALTAVARTARKRKPS PEPEGEVGPPK\TTERPSRGCPHPQRGSRSP*L LHPLLCLRHHPLPHLIPTGPHRLKRPRM\P\SP MAALILVADNAGGSHASKDANQVHSTTRN SNSPPSPSSMNQRRLGPREVGGQGAGNTGGL EPVHPASLPDSSLATSAPLCCTLCHERLEDTH FVQCPSVPSHKFCFPCSRQSIKQQGASGEVYC PSGEKCPLVGSNVPWAFMQGEIATILAGDVK VKKERDS
512	1862	A	3957	1086	3	QDRARLDCSSATSAHCNLRLPGS*DSPASASR VAGTTDTHHHTWLILGSSVQTGFDHVGQAG LELLTSGDPPISASESAGIMGMSHCVWP*SWG LSHHMAPPQGDGGRARGTPGPEQSFWNLSC H*PRCQVPS*LMTQL/FWGRHQYNPTMKRGK LRHREACSLPLPGEGEPGLQPSS*SQNPCSSPLFHHGL*AWLWCPELLLQGQARRH*RSPPS/FK CPATLSLTAWSQTKRLRSQFLLLPWL*RAL*H PP\CHWPSRRSLGDPLLPRSQG*RDGT*ASTFC SYF*DTESHLVAQAGVQWRDLGSLQPPCPRL K\RFSRLSPPSSYTHRYVPSHLAESCISSRDRIP PSRPDRSRNSNSLSR
513	1863	A .	3961	3038	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH VGDHILSIDGTSMEYCTLAEATQFLANTTDQ VKLEILPHHQTRLALKGPDHVKIQRSDRQLT WDSWASNHSSLHTNHHYNTYHPDHCRVPAL TFPKAPPPNSPPALVSSSFSPTSMSAYSLSSLN MGTLPRSLYSTSPRGTMMRRRLKKKDFKSSL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO; in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						SLASSTVGLAGQVVHTETTEVVLTADPVTGF GIQLQGSVFATETLSSPPLISYIEADSPAERCG VLQIGDRVMAINGIPTEDSTFEEASQLLRDSSI TSKVTLEIEFDVAESVIPSSGTFHVKLPKKHN VELGITISSPSSRKPGDPLVISDIKKGSVAHRT GTLELGDKLLAIDNIRLDNCSMEDAVQILQQC EDLVKLKIRKDEDNSDEQESSGAIIYTVELKR YGGPLGVITISGTEEPVFDL*IISSLTKGGLAERT GAIHIGDRILVAINSSSLKGKPLSEAIHLLQMAG ETVTLKIKKQTDAQSASSPKKFPISSHLSDLGD VEEDSSPAQKPGKLSDMYPSHGCPSVDSAVD SWDGSA\IDTSVYGTEGT\SFQASGY\NFNTYD WRSPKQRGSVLSPVT\KPRSQTYPDVGLSYED WDRSTASGFAGAA\DSAETEQEENFWSQALE DLETCGQSGILRELEATIMSGSTMSLNHEAPT PRSPAGSDRPSFQERSSSRPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSPTPVELHKVT LYKDSDMEDFGFSVADGLLEKGVYVKNIRPA GPGDLGGLKPYDRLLQVNHVRTRDFDCCLV VPLIAESGNKLDLVISRNPLASQKSIDQQSLPG D*SEQNSAFFQQPSHGGNLETREPTNTL
514	1864	A	3967	833	800	LEKQGVSGMATKRLARQLGLIRRKSIAPANG NLGRSKSKQLFDYLIVIDFESTCWNDGKHIHH SQEIIEFPAVLLNTSTGQIDSEFQAYVQPQEHPI LSEFCMELTGIKQAQVDEGVPLKICLSQFCK WIHKIQQQKNIIFATGISEPS/DF*SKIMCICYL VR*RISYTY*SKHKSKGC
515	1865	A	3969	492	182	CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL DLLGPEFLARKPLFKTKTYQSTF*SISKNE/FTC PNFIIEEGTDLIF*QVKHNPCHRLTPEEGTVQL NRADS
516	1866	A	3977		1357	KMLC/QKESNYIRLKRAKMDKSMFVKIKTLGI GAFGEVCLARKVDTKALYATKTLRKKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYIPGGDMMSLLIRMGIFPESL ARFYIAELTCAVESVHKMGFIHRDIKPDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGQPPFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV
517	1867	A	3980	1358	1022	FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV
518	1868	A	3986	974	666	SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMGFLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF
-519	1869	A	3994	751	126	NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV HLFSSEMGE\NRPMVG\ARHVYSNAALLSFTP

NO. of No. of no.	ODO TO	CEOTO	Mat	1 650	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
mucleotide seq	SEQ ID	SEQ ID	Met	SEQ ID NO:			
Sequence			1100				F=Phenylalanine G=Glucine H=Victidine
Sequence			ĺ				
Part							M=Methionine N=Asparagine P=Proline
amino acid residue of peptide residue of peptide residue of peptide sequence public sequence T-Thronine, V-Vuine, W-Tryptophan, Y-Tyrosia, E-Vulnow, W-Stop codon, peptide sequence T-Thronine, V-Vuine, W-S		denoc					
Prisidue of peptide Pe	ucnec	}]	'''			T=Threonine, V=Valine, W=Tryptophan.
peptide			ŀ				
		1				50425.500	/=possible nucleotide deletion. \=possible
AHVFADILLITLESYVIPEC			· -	 			LRCLGGEKHKSGLHARPVIVPSLELHYDMDSI
						1	
	520	1870	A	3999	882	698	OSFRLSLLSSWDYRHM*PRLANF*T\FFCRDR/
PPTSASHVAGATGTHHHAWLSV							
1872 1872 A 4015 2 377	521	1871	Α	4011	1346	1178	FFF*ETVSCSAS*AGVRSHDNSSLQPPSPG\SSN
S23							PPTSASHVAGATGTHHHAWLLSV
EYGP-VYSTWSALEGELAPPLEGYSACIONCST	522	1872	A	4015	2	377	QGIALLTRMGESVKHVTGGYKLRTRPLEFAA
AL*ELTDOMTEDFLE*VLREVIL*YSDSMK S23			-		[ĺ	IGDYLDTFALKLGTIDRIAQRIIKEEIEYLVELR
1873		<u> </u>		1	ļ		EYGPVYSTWSALEGELAEPLEGVSACIGNCST
						1	
L*NRREWDEAIKVI,KERQFI-SKMYYPANLSF GNEGDITSPPAK	523	1873	A	4018	341	19	
S24	1		1			:	KVKEVCKTSKS/GQVIYKGVSIRLRANFLAEP
1874		į	ĺ		(
TEPSCLSLPSSWDVRHPPPRLAWLTNELCP+ RQGFTVLARMVLIS*PHDLPASASQSAGITGL SHCSWPTSSILS			1	<u> </u>	<u> </u>		
RQQFTVLARMVLIS*PHDLPASASQSAGITGL	524	1874	A	4020	1067	743	
		ļ	ļ				
1875			1				
PPRLKG/F/SHLSPPSIWDYRRVPCLVNFSIFF			<u> </u>	ļ.,		-	
VETGSCOPCLQLLGSSNPPASASQSAGIAGISH	525	1875	A	4021	781	351	
S26			1			1	
NKIINRPTHPVESSF							
1876		ĺ					
SSAEDGWKADKPYDG*TPGEDHLPTPSFQ	50.6	1056	_	4004	00	241	TROUTER OTER OCCULA A WORD FEVER IN NUM
LHIHISSESQLHHISVKSPPSLSFRLM 527	326	18/6	A	4024	80	341	
1877							
DVAVYFTTKEWAIMG\PAERALYRDVMLEN	527	1077	<u> </u>	4026	502	220	
1878	327	18//	1 ^	4020	393	230	DVAVVETTKEWAIMG\PAFRAI VRDVMI EN
TGSSLSRNDWRAGWIGYLELRRYTYLS							
1878							
VEMGSAKŠVPVTPARPPPHNKHLARVADPRS	528	1878	A	4028	1160	242	<u></u>
PSAGILRTPIQVESSPQPGLPAGEQLEGLKHAQ DSDPRSPTLGIARTEMKTSSGDPPSPLVKQLSE VFETEDSKSNLPPEPVLPPEAPLSSELDLPLGT QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\ NSSKVLGKSPLHPSCQDDNSPGTLTLRQGKA AFKPLSENVSELK\EGA\ LGTGR\LLKTEGRA WEQQQ\HDKENQHFPLVES	320	10/5	1.	1020	1100	- 1-	
DSDPRSPTLGIARTPMKTSSGDPPSPLVKQLSE		1	ì	i			
VFETEDSKSNLPPEPVLPPEAPLSSELDLPLGT QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT ETPVASQSSDKPSRDPETPRSS\GSMRNRWKN NSSKVL\GKSPLHPSCQDDNSPGTLTLRQGKA AFKPLSENVSELK\EGAILGTGR\LLKTEGRA WEQGQD\HDKENQHFPLVES 529 1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*RKIPDITK CW*GCGSTGILIFC\WS*PL*KTI*QPR*FKQI*T ILTIIYSIM*EHTFHNAGV*LSDIYPFMKGYV HTEICT*MFIAVLFVVVKTWKQF 530 1880 A 4057 358 3 LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL RKYGSRINLLKSKHDKNICTENYKT*MKEIEA //DTDKWKDILCSWIRRIHMKDILCSWIGRTHV VKISILPKVNYRFYLISIKIIMAI 531 1881 A 4061 50 278 TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH HYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYFERITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKLI ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR							
ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\ NSSKVL\GKSPLHPSCQDDNSPGTLTLRQGKA AFKPLSENVSELK\GA\ULGTGR\LLKTEGRA WEQQQD\HDKENQHFPLVES S29			1		İ		
ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\ NSSKVL\GKSPLHPSCQDDNSPGTLTLRQGKA AFKPLSENVSELK\GA\ULGTGR\LLKTEGRA WEQQQD\HDKENQHFPLVES S29						ł	QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT
AFKPLSENVSELKÆGA\ILGTGR\LLKTEGRA WEQGQD\HDKENQHFPLVES			Ì				ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\
AFKPLSENVSELKÆGA\ILGTGR\LLKTEGRA WEQGQD\HDKENQHFPLVES			ł	1			
1879 A 4039 2 366 KDMVLIMEMQSMITMKCPQYL*E*RKIPDITK			ļ				
CW*GCGSTGILIFC/WS*PL*KTI*QPR*FKQI*T ILTIIYSIM*EHTFHNAGV*LSDIYPRFMKGYV HTEICT*MFIAVLFVVVKTWKQF 530							
ILTIYSIM*EHTFHNAGV*LSDIYPRFMKGYV HTEICT*MFIAVLFVVVKTWKQF 530 1880 A 4057 358 3 LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL RKYGSRINLLKSKHDKNICTENYKT*MKEIEA //DTDKWKDILCSWIRRIHMKDILCSWIGRTHV VKISILPKVNYRFYLISIKIIMAI 531 1881 A 4061 50 278 TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKLI ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR	529	1879	Α	4039	2	366	KDMVLIMEMQSMITMKCPQYL*E*RKIPDITK
HTEICT*MFIAVLFVVVKTWKQF 530 1880 A 4057 358 3 LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL RKYGSRINLLKSKHDKNICTENYKT*MKEIEA //DTDKWKDILCSWIRRIHMKDILCSWIGRTHV VKISILPKVNYRFYLISIKIIMAI 531 1881 A 4061 50 278 TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKLI ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR			1				CW*GCGSTGILIFC/WS*PL*KTI*QPR*FKQI,*T
530 1880 A 4057 358 3 LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL RKYGSRINLLKSKHDKNICTENYKT*MKEIEA //DTDKWKDILCSWIRRIHMKDILCSWIGRTHV VKISILPKVNYRFYLISIKIIMAI 531 1881 A 4061 50 278 TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYQQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKLITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR	ļ	ł	ł	1	1	1	ILTIIYSIM*EHTFHNAGV*LSDIYPRFMKGYV
RKYGSRINLLKSKHDKNICTENYKT*MKEIEA //DTDKWKDILCSWIRRIHMKDILCSWIGRTHV VKISILPKVNYRFYLISIKIIMAI 531 1881 A 4061 50 278 TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKLI ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR			L	L			
JOTDKWKDILCSWIRRIHMKDILCSWIGRTHV VKISILPKVNYRFYLISIKIIMAI	530	1880	A	4057	358	3	LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL
VKISILPKVNYRFYLISIKIIMAI 531 1881 A 4061 50 278 TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKLI ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR			1			1	RKYGSRINLLKSKHDKNICTENYKT*MKEIEA
531 1881 A 4061 50 278 TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKLI ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR			ł		{	1	/DTDKWKDILCSWIRRIHMKDILCSWIGRTHV
HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKLI ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR			Į _		J	L	
IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKLI ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR	531	1881	Α	4061	50	278	
T*KR 532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKL ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR			l		1	1	HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH
532 1882 A 4069 19 368 NDLLENFKFWE*FKE*LENINGTVTEKETGGV YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKL ITNL/PFIIASKRIKYSGISLTKEMKDLYTETLLR			1	ļ	1		IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF
YKELSSPKYSGTRQFYGQTISNFPGKIISMVY KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKL ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR		<u></u>		L	<u> </u>		
KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL QIWMPVSLMNIVTLKCPT 533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKL ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR	532	1882	Α	4069	19	368	NDLLENFKFWE*FKE*LENINGTVTEKETGGV
QIWMPVSLMNIVTLKCPT		1	1		1		
533 1883 A 4076 1 355 PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKL ITNL/PFIIASKRIKYSGISLTKEMKDLYTETLLR		1	1				KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL
ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR			L		l	l	
	533	1883	A	4076	1	355	PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/
KIKEDTNKWKDI/SCFWVGR/LNIVKMPK/VIC		-	1				ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR
	L	<u> </u>	L	<u> </u>			KIKEDTNKWKDI/SCFWVGR/LNIVKMPK/VIC

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	Ì	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ı	ľ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ŀ	ļ	peptide	'	/=possible nucleotide deletion, \=possible
	1	i	Ì	sequence		nucleotide insertion
						IFNAIPIKMPMMCMAKIEKNSS
534	1884	A	4088	3	1931	IIDSSTRRMESERSPLYRQLIDLGYLSSSHWNC
			1100		.,	GAPGQDTKAQSMLVEQSEKLRHLSTFSHQVL
ł		ł	ļ		}	QTRLVDAAKALNLVHCHCLDIFINQAFDMQR
						DLQITPKRLEYTRKKENELYESLMNIANRKOE
1						EMKDMIVETLNTMKEELLDDATNMEFKDVI
	:					VPENGEPVGTREIKCCIRQIQELIISRLNQAVA
	ļ		[1		NKLISSVDYLRESFVGTLERCLQSLEKSQDVS
	1	ľ		1	ł	VHITSNYLKQILNAAYHVEVTFHSGSSVTRM
						LWEQIKQIIQRITWVSPPAITLEWKRKVAQEAI
ŀ						ESLSASKLAKSICSQFRTRLNSSHEAFAASLRO
ļ	ŀ			}		LEAGHSGRLEKTEDLWLRVRKDHAPRLARLS
]	J	ļ	1			
						LESRSLQDVLLHRKPKLGQELGRGQYGVVYL
	ţ	ŀ				CDNWGGHFPCALKSVVPPDEKHWNDLALEF
	1	ŀ	[HYMRSLPKHERLVDLHGSVIDYNYGGGSSIA
•	1			ĺ	ĺ	VLLIMERLHRDLYTGLKAGLTLETRLQIALDV
l .		ŀ				VEGIRFLHSQGLVHRDIKLKNVLLDKQNRAKI
ľ						TDLGFCKPEAMMSGSIVGTPIHMAPELFTGK
				1		YDNSVDVYAFGILFWYICSGSVKLPEAFERCA
ļ		j	j	j	ļ	SKDHLWNNVRRGARPERLPVFDEECWQLME
						ACWDGDPLKRPLLGIVQPMLQGIMNRLCKS\
525	1005		4000			NSEQPNRGLDDST
535	1885	Α	4090	2	417	ALMPHEANYEEIFLKTDKDMDGFESGLEVRE
]	Ì			IFLKTR/GLPSTLLAHIWALCDSKDCGKLSKD
i	1				ŀ	HFALAFHLIT\QKLIKGIDPPLVLTPEKISPSNR
						ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK
	1006					HNRKRIWLRA
536	1886	Α	4102	569	829	DHQK*KNIPCSWIGRINIVKMSILPKAIYRFSAI
	Ì					PIKIPMTFFTEI*S*NVYRTTKTQE*AKAILSKK
<u> </u>		<u></u>				EQNLEESHYLDFK*YYRAV
537	1887	Α	4104	54	281	SIDCEHLIRRMLVLDPSKRLTIAQIKEHKWML
		ļ				IEVPVQRPVLYPQEQENEPSIGEFNEQVLRLM
						HSLGIDQQKTIE
538	1888	Α	4109	141	314	IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL
						VPRENITAWMNAIGLIITALPVS
539	1889	A	4111	268	1	ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM
						I*KFAT\KKSPGPYRFTAEFSHTFKEDLVPILW
						PLFPKIYREGTLPHSFYEASITL
540	1890	A	4142	198	2064	PEPGAGRAATPWGPLFWRGRGSGRCEKAAE
		1	į			AALGDFLGLHRRTQQPAVDRLLSDASAQWR
			1			VRGHGGVRESGRAPQQPGRRRGRRPRKRPR
]	1]		GRWRREGCGAGGRGVCVAAWSQRSIAGNN
	1				l	DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH
[1]	1	[IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR
l .	1	1	1		1	CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT
· .		l	1	ĺ	ĺ	MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY
]				GFPELEDSTSEYLCTILNIQNVCMTFDVASLY
		1	ł			SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK
	1					TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK
1]	<u> </u>		j		ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP
				ļ		DAILDAIKVRSESRDMDLNYRGMLIPEENIAT
!			ļ			MKYGAQVVKGELKSALLDGDTQNYDLDHG
1		1	1			FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR
	1		1	[1	DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS
}	1	}	ļ]		WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF
		1			_	ECMFTNKTFILEKGLIVPMENVATIADCASVI
		1	1	I		EGVSRSRNALLNGDTKNYDWDSGYTCHQLG
541	1891	<u> </u>	4145	202	770	SGAIVVQLAQPYMIGSIRVLLWDCDDRSY
1 271	1071	A	4146	282	778	GTLGYPNGARGQPQDNFFAHQ\VSHHPPISAC

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	ł	l	i	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
İ		i		peptide		/=possible nucleotide deletion, \=possible
		ì		sequence		nucleotide insertion
						HAESENFAFWQDMKWKNKFWGKSLEIVPVG
ł		1				TVNVSLPRFGDHFEWNKVTSCIHNVLSGQRW
1		1				IEHYGEVLIRNTQDSSCHCKITFCKAKYWSSN
1	1	l				VHEVQGAVLSRSGRVLHRLFGKWHEGLYRG
1	1	ĺ	l			PTPGGQCIWKP
542	1892	A	4147	44	433	SVDAYVCNDIVFSYRTTITLLEGA*LTHRYVA
342	1652	1	4147		433	
	l					QDPKQGQLRSLHLTCDSAPAGSQGTWSTSCR
	Ì	l	ļ			NHLIFRGGAQITFLATFDDSPKAVLGDRLLLT
]		ì			ANVSSENNTPRTSKTTFQLELSVKDAVYTVV
L						SSH
543	1893	Α	4153	678	11	TISYPOCLTOMYFLISFANVDTFLLPIMALDH
		1]	YVAICSALQ*CSITTP/ELCQGLPVLA*AGSSLIS
1	1			1	į	PVHTVIMSRLAFCSSAQISHFYRDAYLLMKIA
1		J]		j	CSHT*\NQHVFLGAVVLFLAPCALILVSYIRIA
1	1	}	1			AAILRIPSPTRRRKACSICSSHLSLVTLFYGTV
						LGICI*PPDSFSAQDAIATIMYTVVTSMLNPFIY
	j		•			SLMNKEVQEAVRRLFSRGSHSSWCW
544	1894	A	4158	3	538	LLYAQAGVQ*LNLSSLQPQPAGLKQSSHPSLP
777	1054	1	1 4150	,	330	SSWDYRYSTPHPANFFVEMEFHHVAQAGLEL
1	1					LGSGDLPTSTSHSAGITGV\SHHAPPRLISSEGS
1	ì		ł			
	1		į			LLGHLLCLPMVFPLLCVFVLISSSLAGEEAAG
			į.			LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE
						LKVGREGHVLPWQAHVVEF
545	1895	Α	4160	1	412	HPLGLGLVPSEIFSPQDKKAADGSILAPARGE
						DLEAGLKGSFMDGRLQASVSVFRIQRVGSAM
1						QDTASAMPCLPYYPTSHCFMAGGKSRSQGW
	1	ļ				ELELSGEPAPGWQVLAGYTYTQARYLRDASE
		1	1			ANVGQPLRPVDPR
546	1896	A	4174	1252	1190	FFQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL
	1		ŀ			OPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV
						FLVETGFLHV\GQ\ASLELPTSGDTPAS\ASQSA
	j .	1	1			GITGVSHHA*PRASGRRCW
547	1897	A	4176	3029	1	AGPDGLAAPASCQGARGQTRVPGAFSWLAP
1 347	1007	^	1170	3027	1	GSHHASEGLAPGVPPAGGVSAQELTAPPQEG
1						WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA
	1	ļ		Ī		
1	l	}	l	ł	ł	RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS
1	,			1		GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ
1	i		1			PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL
	1	1]		İ	AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP
	1	I	1		1	HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS
ſ	í	1	ſ		1	TMAPIPSALAVWEPAGSSPQLSSAPADSS\PLP
	l		ļ	l]	ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/
			1			RCPPACSPAAASSFSFESQPCPSAPSKASPAPA
	1		i			AL/IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL
1	ł	Į	}	l	ļ	AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA
1 .	ŀ	1]			LAS\PLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/
					ŀ	PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP
	1		1			AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ
				1	1	ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP
1	ł	1	ł	}	1	PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP
	1	i	[1	1	RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC
1		ł	1			
		1	1			LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS
1	1	ŀ	1		1	WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL
	1	1			ĺ	PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ
1	1	·	1			PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL
1		1			1	TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP
-	1	1	1]	MP*CFHRPSPPLP/LSSPFPA/LRPQAPQFPLHLP
ł	J		J	[}	P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT
1		Į.			i	PPLWPSLALLP*PLPPPPPVPSFSASLLCSLPAH

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTPASPGLGRSCLGKPQTLPWISFWPPSGRLA PGTWQPW/PVSPAPLSCLSAWDPWELPSPQPQ VCSTAELPTSCLLSSPGP\PAFQPPRFGCL*GPP GPPGLPPLQSSLSFPPPPPPVPQPPAPPALQWG LHLPGGRTK
548	1898	A	4180	2369	844	RIHREEDFQFILKGIARLLSNPLLQTYLPNSTK KIQFHQELLVLFWKLCDFNKVGQPRGALQGD GEQLPQ*PGGRDSVRLRGVGQSCPSLELSPLG PSPHP*KFLFFVLKSSDVLDILVPILFFLNDAR ADQSRVGLMHIGVFILLLLSGECNFGVRLNKP YSIRVPMDIPVFTGTHADLLIVVFHKIITSGHQ RLQPLFDCLLTIVVNVSPYLKSLSMVTANKLL HLLEAFSTTWFLFSAAQNHHLVFFLLEVFNNI IQYQFDGNSNLVYAIIRKRSIFHQLANLPTDPP TIHKALQRRRRTPEPLSRTGSQGGAPPWRAPA PLPLQSQAPSRPVWWLLQALTS*PRSPRCQR MAPCGPWNLSPSRAWRMAARLRGSPARHGG SSGDRP/HSSASGQWSPTPEWVLSWKSKLPLQ TIMRLLQVLVPQVEKICIDKGLTDESEILRFLQ HGTLVGLLPVPHPILIRKYQANSGTAMWFRT YMWGVIYLRNVDPPVWYDTDVKLFEIQRV
549	1899	Α	4191	858	321	LPWQRLGVLLSRGKMAVTGWLESLRTAQKT ALLQDGRRKVHYLFPDGKEMAEEYDEKTSE LLVRKWRVKSALGAMGQWQLEVGDPAPLG AGNLGPELIKESNANPIFMRKDTKMSFQWRIR NLPYPKDVYSVSVDQKERCIIVRTTNKKYYK KFSIPDLDRHQLPLDDALLSFA\TPTAP
550	1900	A	4192	1	1980	IRHTGSDIAGVCGWLLLSGPCGVGLDLDSRLL GASAMRRSEVLAEESIVCLQKALNHLREIWE LIGIPEDQRLQRTEVVKKHIKELLDMMIAEEE SLKERLIKSISVCQKELNTLCSELHVEPFQEEG ETTILQLEKDLRTQVELMRKQKKERKQELKL LQEQDQELCEILCMPHYDIDSASVPSLEELNQ FRQHVTTLRETKASRREEF/VSSIKRQIILCME ELDHTPDTSFERDVVCEDEDAFCLSLENIATL QKLLRQULEMQKSQNEAVCEGLRTQIRELW DRLQIPEEEREAVATIMSGSKAKVRK\ALQULE VDRLEELEKCKTMKKVIEAIRVELVQYWDQC FYSQEQRQAFAPFCAEDYTESLLQLHDAEIVR LKNYYEVHKELFEGVQKWEETWRLFLEFER KASDPNRFTNRGGNLLKEEKQRAKLQKMLP KLEEELKARIELWEQEHSKAFMVNGQKFME YVAEQWEHRLEKERAKQERQLKNKKQTET EMLYGSAPRTPSKRRGLAPNTPGKARKLNTT TMSNATANSSIRPIFGGTVYHSPVSRLPPSGSK PVAASTCSGKKTPRTGRHGANKENLELNGSI LSGGYPGSAPLQRNFSINSVASTYSEFADPSLS DSSTVGLQRELSKASKSDATSGILNSTNIQS
551	1901	Α	4194	3	1008	AWHEGLVSSPAIGAYLSASYGDSLVVLVATV VALLDICFILVAVPESLPEKMRPVSWGAQISW KQADPFASLKKVGKDSTVLL\ICITVCLSYLPE AG\QYSSFF\LYLR\QVIGFG\TVKIAAFIAMVGI LSIVAQTAFLSILMRSLGNKNTVLLGLGFQML QLAWYGFGSQAWMMWAAGTVAAMSSITFP AISALVSRNAESDQQGVAQGIITGIRGLCNGL GPALYGFIFYMFHVELTELGPKLNSNNVPLQ GAVIPGPPFLFGACIVLMSFLVALFIPEYSKAS GVQKHSNSSSGSLTNTPERGSDEDIEPLLQDS SIWELSSFEEPGNQCTEL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in NO:	nucleotide	location	
eotide	seq-	1	USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	
uence	40.50		914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
Lonce		ł	'''	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ĺ	ĺ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
}				sequence		nucleotide insertion
552	1902	A	4197	2	14302	ARPPPAPGSRQQKQKAAPGAAAAAELRGAR
352	1702	А	4177	2	14302	EPAPARRGTMADGGEGEDEIOFL'RTDDEVV
	ļ					LQCTATIHKEQQKLCLAAEGFGNRLCFLESTS
1			i			NSKNVPPDLSICTFVLEQSLSVRALQEMLANT
}						VEKSEGOVDVEKWKFMMKTAOGGGHRTLL
						YGHAILLRHSYSGMYLCCLSTSRSSTDKLAFD
1						VGLQEDTTGEACWWTIHPASKQRSEGEKVR
						VGDDLILVSVSSERYLHLSYGNGSLHVDAAF
[]			1			QQTLWSVAPISSGSEAAQGYLIGGDVLRLLH
						GHMDECLTVPSGEHGEEQRRTVHYEGGAVS
						VHARSLWRLETLRVAWSGSHIRWGQPFRLR
i l						HVTTGKYLSLMEDKNLLLMDKEKADVKSTA
1						FTFRSSKEKLDVGVRKEVDGMGTSEIKYGDS
						VCYIQHVDTGLWLTYQSVDVKSVRMGSIQR
i						KAIMHHEGHMDDGISLSRSQHEESRTARVIRS
						TVFLFNRFIRGLDALSKKAKASTVDLPIESVSL
						SLQDLIGYFHPPDEHLEHEDKQNRLRALKNR
1						QNLFQEEGMINLVLECIDRLHVYSSAAHFAD
J				-		VAGREAGESWKSILNSLYELLAALIRGNRKN
					•	CAQFSGSLDWLISRLERLEASSGILEVLHCVL
						VESPEALNIIKEGHIKSIISLLDKHGRNHKVLD
						VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL
						LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY
i i						YELMVDHTEPFVTAEATHLRVGWASTEGYSP
						YPGGGEEWGGNGVGDDLFSYGFDGLHLWSG
[CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF
1						RINGQPVQGMFENFNIDGLFFPVVSFSAGIKV
						RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL
1 1				ĺ	•	KVEHSREYKQERTYTRDLLGPTVSLTQAAFT
]						PIPVDTSQIVLPPHLERIREKLAENIHELWVMN
		1				KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ
						ERNYNLQMSLETLKTLLALGCHVGISDEHAE
	}	1				DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT
						PSQEAMVDKLAENAHNVWARDRIRQGWTY
		1				GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS
]]		J		j	J	LREAVRTLLGYGYNLEAPDQDHAARAEVCS
					ļ	GTGERFRIFRAEKTYAVKAGRWYFEFETVTA
					ŀ	GDMRVGWSRPGCQPDQELGSDERAFAFDGF
		-			1	KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD
					İ	GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC
	1	ł		}	ł	GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV
		ļ			ł	PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN
]		SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG
	1	1		ļ	l	LFGPKNDLEDYDADSDFEVLMKTAHGHLVP
ļ		l	1		ŀ	DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ
	1	i	Į		ļ	RFLLRRTKPDYSTSHSARLTEDVLADDRDDY
1 1		ļ	ļ	Ì	ļ	DFLMQTSTYYYSVRIFPGQEPANVWVGWITS
		1				DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE
[-				SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC
j	1		1	1	ļ	VVDAASGLLTFIANGKELSTYYQVEPSTKLFP
		1	1	1		AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS
	1	ì	- 1	ł		EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL
	i		ļ			KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN
]		1	ļ			RSVDILELTEQEELLKFHYHTLRLYSAVCALG
	1	ł	ł	ŀ		NHRVAHALCSHVDEPQLLYAIENKYMPGLLR
	.	İ	l	l	İ	AGYYDLLIDIHLSSYATARLMMNNEYIVPMT
[1		ŀ	ļ		EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF
ļ		ł	l	1		SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE
	İ			l		AVKEGSLHARDPVGGTTEFLFVPLIKLFYTLLI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
}				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
				sequence		MGIFHNEDLKHILQLIEPSVFKEAATPEEESDT
						LEKELSVDDAKLQGAGEEEAKGGKRPKEGLL
						QMKLPEPVKLQMCLLLQYLCDCQVRHRIEAI VAFSDDFVAKLQDNQRFRYNEVMQALNMSA
		•				ALTARKTKEFRSPPQEQINMLLNFKDDKSECP
ļ						CPEEIRDQLLDFHEDLMTHCGIELDEDGSLDG NSDLTIRGRLLSLVEKVTYLKKKQAEKPVES
ļ						DSKKSSTLQQLISETMVRWAQESVIEDPELVR
						AMFVLLHRQYDGIGGLVRALPKTYTINGVSV EDTINLLASLGQIRSLLSVRMGKEEEKLMIRG
						LGDIMNNKVFYQHPNLMRALGMHETVMEV
						MVNVLGGGESKEITFPKMVANCCRFLCYFCR ISRONOKAMFDHLSYLLENSSVGLASPAMRG
						STPLDVAAASVMDNNELALALREPDLEKVVR
						YLAGCGLQSCQMLVSKGYPDIGWNPVEGER YLDFLRFAVFCNGESVEENANVVVRLLIRRPE
						CFGPALRGEGGNGLLAAMEEAIKIAEDPSRD
į	}				·	GPSPNSGSSKTLDTEEEEDDTIHMGNAIMTFY SALIDLLGRCAPEMHLIHAGKGEAIRIRSILRS
1						LIPLGDLVGVISIAFQMPTIAKDGNVVEPDMS
						AGFCPDHKAAMVLFLDRVYGIEVQDFLLHLL EVGFLPDLRAAASLDTAALSATDMALALNRY
1	:					LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT
						VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS MMQHLLRRLVFDVPLLNEHAKMPLKLLTNH
ľ	}				•	YERCWKYYCLPGGWGNFGAASEEELHLSRK
į						LFWGIFDALSQKKYEQELFKLALPCLSAVAG ALPPDYMESNYVSMMEKQSSMDSEGNFNPO
						PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK
}						LANGWIYGEIYSDSSKVQPLMKPYKLLSEKE KEIYRWPIKESLKTMLARTMRTERTREGDSM
ļ						ALYNRTRISQTSQVSVDAAHGYSPRAIDMS
						NVTLSRDLHAMAEMMAENYHNIWAKKKKM ELESKGGGNHPLLVPYDTLTAKEKAKDREKA
						QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA
						YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF PYEOEIKFFAKVVLPLIDOYFKNHRLYFLSAA
į						SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH
1						RISLFGNDATSIVNCLHILGQTLDARTVMKTG LESVKSALRAFLDNAAEDLEKTMENLKQGQF
						THTRNQPKGVTQIINYTTVALLPMLSSLFEHI
	}					GQHQFGEDLILEDVQVSCYRILTSLYALGTSK SIYVERQRSALGECLAAFAGAFPVAFLETHLD
ļ						KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP
						SLEKLMEEIVELAESGIRYTQMPHVMEVILPM
1						LCSYMSRWWEHGPENNPERAEMCCTALNSE HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF
						SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV
						SEEDHLKAEARGDMSEAELLILDEFTTLARDL YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR
						MVAEVFIYWSKSHNFKREEONFVVONEINN
						MSFLITDTKSKMSKAAVSDQERKKMKRKGD RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA
						LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP
						AIRWQMALYKDLPNRTDDTSDPEKTVERVL DIANVLFHLEQKSKRVGRRHYCLVEHPORSK
					- 4	KAVWHKLLSKQRKRAVVACFRMAPLYNLPR
	ļ					HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT
L			<u> </u>			EKCKLEEDFLYMAYADIMAKSCHDEEDDDG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
ĺ	<u> </u>		{	peptide sequence		nucleotide insertion
						EEEVKSFEEKEMEKQKLLYQQARLHDRGAA EMVLQTISASKGETGPMVAATLKLGIAILNGG NSTVQQKMLDYLKEKKDVGFFQSLAGLMQS CSVLDLNAFERQNKAEGLGMVTEEGSGEKV LQDDEFTCDLFRFLQLLCEGHNSDFQNYLRT QTGNNTTVNIIISTVDYLLRVQESISDFYWYY SGKDVIDEQGQRNFSKAIQVAKQVFNTLTEYI
						QGPCTGNQQSLAHSRLWDAVVGFLHVFAHM QMKLSQDSSQIELLKELMDLQKDMVVMLLS MLEGNVVNGTIGKQMVDMLVESSNNVEMIL KFFDMFLKLKDLTSSDTFKEYDPDGKGVISK RDFHKAMESHKHYTQSETEFLLSCAETDENE TLDYEEFVKRFHEPAKDIGFNVAVLLTNLSEH
						MPNDTRLQTFLELAESVLNYFQPFLGRIEIMG SAKRIERVYFEISESSRTQWEKPQVKESKRQFI FDVVNEGGEKEKMELFVNFCEDTIFEMQLAA QISESDLNERSANKEESEKERPEEQGPRMAFF SILTVRSALFALRYNILTLMRMLSLKSLKKQM KKVKKMTVKDMVTAFFSSYWSIFMTLLHFV
						ASVFRGFFRIICSLLLGGSLVEGAKKIKVAELL ANMPDPTQDEVRGDGEEGERKPLEAALPSED LTDLKELTEESDLLSDIFGLDLKREGGQYKLIP HNPNAGLSDLMSNPVPMPEVQEKFQEQKAK EEEKEEKEETKSEPEKAEGEDGEKEEKAKED KGKQKLRQLHTHRYGEPEVPESAFWKKIIAY
						QQKLLNYFARNFYNMRMLALFVAFAINFILL FYKVSTSSVVEGKELPTRSSSENAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVRILPILHTVISF FCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKKPKKDSSLSAVLNSIDVKYQMW
		ļ				KLGVVFTDNSFLYLAWYMT
553	1903	A	4199	31	767	LPELNGRGAGLRRAEPSERGGGAERTQQVAA LPLSHGHSHGGGGCRCAAER/VGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERAD\DRSKFVESDADEELLFNIPFTG\HVKLK GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS
554	1904	A	4200	1	961	GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGT\ RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH
555	1905	A	4211	331	2419	KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNP NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPLS DSNRDHTANRQQRSTSPVARRTRSQTSVNFN GSSSNIPRTRLASRGQNPAEGSFSTLGRLRNGI

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, E=Phonylologica, C=Clusica, Hallington
nucl- eotide	peptide seq-		in USSN	nucleotide location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		İ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ľ			peptide		/=possible nucleotide deletion, \=possible nucleotide insertion
	 -	 -	<u> </u>	sequence	<u> </u>	GGAAGIPRANASRTNFSSHTNQSGGSELRQRE
						GQRFGAAHVWENGARSNVTVRNTNQRLEPI RLRSTSNSRSRSPIQRQSGTVYHNSQRESRPV
						QQTTRRSVRRRGRTRVFLEQDRERERRGTAY
						TPFSNSRLVSRITVEEGEESSRSSTAVRRHPTIT
[[[[LDLQVR\RIRPGENRDRDSIANRTRSRVGLAE
						NTVTIESNSGGFRRTISRLERSGIRTYVSTITVP LRRISENELVEPSSVALRSILRQIMTGFGELSSL
						MEADSESELQRIGOHLPDMHSELSILGTDN
	Ì	İ	ļ			NRSQHREGSSQDRQAQGDSTEMHGENETTQP
1						HTRNSDSRGGRQLRNPNNLVETGTLPILRLAH
		1				FFLLNESDDDDRIRGLTKEQIDNLSTRHYEHN
						SIDSELGKICSVCISDYVTGNKLRQLPCMHEF
556	1906	A	4212	3	462	HIHCIDRWLSENCTCPICRQPVLGSNIANNG LQRQRQHPAAAPAVPVRCFTFCFTDIVIMPKR
	1700	``			102	KSPENTEGKDGSKVTKQEPTRRSARLSAKPA
						PPKPEPKPRKTSAKKEPGAKISRGAKGKKEEK
]					QEAGKEGTAPSENGETKAEEIHISRSTVNVST
557	1907	A	4213	774	507	SRGTPPSTLSVKGQIETVRVKGTEN
337	1907	A	4213	774	507	ARRFSCLTLQTSWGHRH\GPPRP\ANFVFLVET GFLHIGQAGHKLPTSGDPPASASOSARITGMS
				ļ		HRTWFLASFLIDSCKNFIVYKIMYTL
558	1908	A	4225	3	1253	TYRHAEREHPETSSATKVSYDYRHKRPKLLD
1						GDQDFSDGRTQKYCKEEDRKYSFQKGPLNRE
ł	ł					LDCFNTGRGRETQDGQVKEPFKPSKKDSIAC
		1	ļ			TYSNKNDVDLRSSNDKWKEKKKKEGDCRKE SNSSSNQLDKSQKLPDVKPSPINLRKKSLTVK
		Ì				VDVKKTVDTFRVASSYSTERQMSHDLVAVG
						RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT
1			1			IIHQVKANYFPSPGITLHERFS\KMADIHKADV
		-				NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE
						QTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI
						ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ
-						KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK
						KKVP
559	1909	Α	4235	1	323	KFSIPFFLRWSFTLV\PRLEGNDMISVHCNLGL
	l					LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS
		[1			HHAWLQVLNS
560	1910	A	4246	2	1569	TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ
		ļ				AALVNYSRLSEYAKIEGKKREMYELPVFCLA
1	ŀ					SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR
]			LAELVIEVLQQNEEHHAEAFAWWSDLMVEH
						AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LL\NDFLRTGLLICGNGK\FHKHLQDLFAPLVV
						R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN
	!	1				GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK
}	1	1	1	l .		HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK
		1				TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP
						KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS
}]	<u> </u>				SFLSFTVKAASKYVDVPKPGMDVADAYVTF
1		ľ	{		1	VRHSQDVLRDKVNEEMYIERLFDQWYNSSM
						NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY
		1				RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA
561	10:	<u> </u>	4252	1200	-	SVSEGGGLQGISMKDSDEEDEEDD
561	1911	A	4257	1300	654	SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI FPDLFKRAAERLQYVFGYKLVELEPKSNTYIL
L	L	L	<u>[</u>	L	L	II DELIGIONE LE VEUE LE PRONTYIL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion INTLEPVEEDAEMRGDQGTPTTGLLMIVLGLI FMKGNTIKETEAWDFLLAL\GVYPTKKHLIFG DPKKLITEDFVRQRYLEYRRIPHTDPVDYEFQ WGPRTNLETSKMKVLKFVAKVHNQDPKDW
562	1912	A	4260	1	1498	PAQYCEALADEENRARPQPSGPAPSS MVTWLYRFLPTSNMAAKLRSLLPPDLRLQF WLHARLQKCFLSRGCGSYCAGAKASPLPGK MAMGLMCGRRELLRLLQSGRRVHSVAGPSQ WLGKPLTTRLLFPAAPCCCRPHYLFLAASGPR SLSTSAISFAEVQVQAPPVVAATPSPTAVPEV ASGETADVVQTAAEQSFAELGLGSYTPVGLI QNLLEFMHVDLGLPWWGAIAACTVFARCLIF PLIVTGQREAARIHNHLPEIQKFSSRIREAKLA GDHIEYYKASSEMALYQKKHGIKLYKPLILPV TQAPIFISFFIALREMANLPVPSLQTGGLWWF QDLTVSDPIYILPLAVTATMWAVLELGAETG VQSSDLQWMRNVIRMMPLITLPITMHFPTAV FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQR VVHDLDKLPPREGFLESFKKGWKNAEMTRQ LREREQRMRNQLELAARGPLRQTFTHNPLLQ PGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG
563	1913	A	4265	623	116	MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQN THGTRRTLSSGPSNNLPLPLSSSATMPSMQCK HRSPNGGLFRQSPVK/TPPIPMSFQPVPGGV\L PRGSGNPPHGTSILTAPPALLPHPPTHPTQQSF LIQENNNTNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF
564	1914	A	4270	3	368	ILKRKLSSLNSEVSTIQNTRMLAFKATAQLFIL GCTWCLGLLQVGPAAQVMAYLFTIINSLQGF FIFLVYCLLS\QQVQKQYQKWFREIVKSKSES ETYTLSSKMGPDSKPSEGDVFPRTSE
565	1915	A	4288	83	406	RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS PPSALLAPTKPRALGTLRLYECSPELCTTMLP PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP GQTGASRTPRT
566	1916	A	4298	1041	229	LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ LLKKNGGIVMVTLSMGVLQCNLLANVSTVA DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGL\E DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH FHLGASEWTPRLLIWR
567	1917	A	4299	1	1106	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD INEAYVETLKHCFMMPQSLGVIGGKPNSAHY FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF GAECCLGMTRKTFGFLRFFFSMLG
568	1918	A	4300	2012	1843	SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS LMSVLIPKLPQLHGVRIFGINKY
569	1919	A	4302	186	531	WTFCLFL/WWVPESARWLLTQGHVKEAHRY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLHCARLNGRPVCEDSFSQEVRVNVCVSMHI CVWWGVGCVKCLPPRAHHIWQEKPLGPHRT
-500	1000		4000			VTESKLEAEGKTKEKAREKERKKKS
570	1920	Α	4308	3		RSGQGKVYGLIGRRRFQQMDVLEGLNLLITIS GKRNKLRVYYLSWLRNKILHNDPEVEKKQG WTTVGDMEGCGHYRVVKYERIKFLVIALKSS VEVYAWAPKPYHKFMAFKSFADLPHRPLLV DLTVEEGQRLKVIYGSSAGFHAVDVDSGNSY DIYIPVHIQSQITPHAIIFLPNTDGMEMLLCYE DEGVYVNTYGRIIKDVVLQWGEMPTSVAYIC SNQIMGWGEKAIEIRSVETGHLDGVFMHKRA QRLKFLCERNDKVFFASVRSGGSSQVYFMTL NRNCIMNW
571	1921	A	4309	9	524	ASREMDVTKVCGEMRYQLNKTNMEKDEAE KEHREFRAKTNRDLEIKDQEIEKLRIELDESK QHLEQEQQKAALAREECLRLTELLGESEHQL HLTRQEKDSIQQSFSKEAKAQALQAQQREQE LTQKIQQMEAQHDKTENEQYLLLTSQNTFLT KLKEECCTLAKKLEQISQ
572	1922	A	4318	1	1119	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGL\T DINEAYV\ETL\KHCFHGWPQFPG/VVHREGK PNSAHYFIGYVGEELIYLDPHTTQPAVEPTDG CFIPDESFHCQHPPCRMSIAELDPSIAVVRGGH LSTQAFGAECCLGMTRKTFGFLRFFFSMLG
573	1923	A	4333	363	1066	GGVPVGLASKPFQILYGHTNEVLSVGISTELD MAVSGSRDGTVIIHTIQKGQYMRTLRPPCESS LFLTIPNLAISWEGHIVVYSSTEEKTTLKERM HYICFSINGKYLGSQILKEQVSDICIIGEHIVTG SIQGFLSIRDLHSLNLSINPLAMRLPIHCVCVT KEYSHILVGLEDGKLIVVGVGKPAEVKPSISN FISHAVGDYFGSPSFQLIEKSPLGINKLKAKFD FSKGSK
574	1924	A	4346	359	1234	MDTLEEVTWANGSTALPPPLAPNISVPHRCLL LLYEDIGTSRVRYWDLLLLIPNVLFLIFLLWK LPSARAKIRITSSPIFITFYILVFVVALVGIARA VVSMTVSTSNAATVADKILWEITRFFLLAIEL SVIILGLAFGHLESKSSIKRVLAITTVLSLAYSV TQGTLEILYPDAHLSAEDFNIYGHGGRQFWL VSSCFFFLVYSLVVILPKTPLKERISLPSRRSFY VYAGILALLNLLQGLGSVLLCFDIIEGLCCVD ATTFLYFSFFAPLIYVAFLRGFFGSEPKILF
575	1925	A	4360	2038	1512	GCWWRHPWLASQRDCLDCRIQLAEKFVKAV SKPSRPDMNPIRVKEVYRLEEMEKIFVRLEM KIIKGSSGTPKLSYTGRDDRHFVPMGLYIVRT VNEPWTMGFSKSFKKKFFYNKKTKDSTFDLP ADSIAPFHICYYGRLFWEWGDGIRVHDSQKP QDQDKLSKEDVLSFIQMHRA
576	1926	A	4365	69	500	QVEGRQGREVKRTAWRISPVWRPARCRRRST PQP/PE/PGAQQQERHRQGEAPMQALDPRAEP GPQAQSHAACQPEPEPPRVLLDPTAARGGVQ GRP/GLSRHPGLAPHPQTHTPWPQSGRLPCAS EPLPLGGIRPTPGLEPKGRDLM

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	""	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ł	l	peptide		/=possible nucleotide deletion, \=possible
<u></u>				sequence		nucleotide insertion
577	1927	A	4366	785	502	SAPPKKKNGVLFLSPRLKSSGAIWVHSTPTLW
İ						ASSNSRASTPKVAGITGARPHARIIFVFLIEMG
570	1000		4267		201	FHNVGQAGL/DTLTLVICPPQPPKLLGLQM
578	1928	A	4367	1	221	FFFFLKKSRCVTQAGVQG\PISLHPPPPGFKRF SRLSLLSSWDYRHP/HAANFCIFSRDG\VSPYW
i		ļ				SGWSRTPDLR
579	1929	Ā	4383	1	224	FETESHSVTQAGMQWHNLGSLQPMP/PGLKR
3/3	1929	1 ^	4303	1	224	FSCLRLQSSWDHRHAPPHLAHFCIFSRDGVSP
						CWPGWSSTPDLK
580	1930	A	4397	410	94	SRLKPYSTNVTAKKLPATNIPNLDCFTAKLYQ
500	1320	l			*	\VFKKGI\IHILHELFQNKEEGAFPNS/FYEASFT
	,	İ				LRPKSDRDIAKEESYSTISLLSTDTKILMSKYK
						QLKSSDL
581	1931	A	4414	670	3	VLVHRQCGGILRLRRKEAVSVLDSADIEVTDS
						RLPHATIVDHRPQHRWLETCNAPPQLIQGKA
						RSAPKPSQASGHFSVELVRGYAGFGLTLGGG
	[RDVAGDTPLAVRGLLKDGP\AQRCGRLEVGD
						LVLHINGESTQGLT\HAQAVERIRAGGPQLHL
	[}	VIRRPLETHPGKPRGVGEPRKGVVPSWPDRSP
						DPGGPEVTGSRSSSTSLVQHPPSRTTLKKTRG
582	1932	A	4424	194	449	SPE VLYIRKKKRLEKLRHQLMPMYNFDPTEEQDE
382	1932	A	4424	194	449	LEQELLEHGRDAASVQAATSVQAMQGKTTL
	1			·		PS\QGPLQRPSRLVFT\DVANAIHV
583	1933	A	4435	1	166	APGPPVPPPGSPPEQMPGPCPASMPP/DPPPGS
100	1,000	**	'''	-		PPEQMPGPCPVSAPP/GPPPGSPPEQMPGPCPV
	ļ					SAPPALLQDTSV
584	1934	A	4439	1	628	SATPQQPSAPQHQGTLNQPPVPGMDESMSYQ
						APPQQLPSAQPPQPSNPPHGAHTLNSGPQPGT
			İ			APATQHSQAGPATGQAYGPHTYTEPAKPKK
						GQQLWNRMKPAPGT\EVSSSTSRSDPLLLPPR
		ŀ	Ì			ALAPTQRASTVVLAPSPT/SEKVQNHSGSSAR
	1	l			ļ	GNLSGKPDDWP/LGHERVCGALLHRL*VGGG
505	1935		4463	10	144	QGPHGKAAQGGAAGAAAGRLGLYH HKPVTNSRDTQEVPLEKAKQVLKIIATFKHTT
585	1935	Α	4403	10	144	SIFDDFAHYEKRQ
586	1936	Α	4464	1309	103	LNAESYVSFTTKLDIPTAAKYEYGVPLQTSDS
300	1,50	A		1505	103	FLRFPSSLTSSLCTDNNPAAFLVNQAVKCTRK
			1		1	INLEQCEEIEALSMAFYSSPEILRVPDSRKKVPI
		1				TVQSIVIQSLNKTLTRREDTDVLQPTLVNAGH
'		1				FSLCVNVVLEVKYSLTYTDAGEVTKADLSFV
						LGTVSSVVVPLQQKFEIHFLQENTQPVPLSGN
	1	l	1		1	PGYVVGLPLAAGFQPHKGSGIIQTTNRYGQLT
1		}	ĺ			ILHSTTEQDCLALEGVRTPVLFGYTMQSGCK
	}	Ι.				LRLTGALPCQLVAQKVKSLLWGQGFPDYVA
			1			PFGNSQGP/ADMLDWVPIHFITQSFNRKDSCQ
						LPGALVIEVKWTKYGSLLNPQAKIVNVTANLI
		1				SSSFPEANSGNERTILISTAVTFVDVSAPAEAG
507	1027		4471	614	207	FRAPPAINARLPFNFFFFFV
587	1937	A	4471	614	387	LLGRASAC/LQLQSSW/D/HRPMLPYLANFVF CKDR/SFTWLPRLVLNSWLQVILLPWPPTGCD
		1				NKHEPPCPATKRRHSGSI
588	1938	A	4480	1720	1458	HDLGSLQPPPPGFKRFSCLSLPSSWDYRLMPP
700	1536	^	7700	1,20	1450	CPANFCIIII/DFLVETGFHHVGQASHELLTSGD
1						PPTSASQSAGITGMSYHTWFGES
589	1939	A	4487	922	332	APVTTSPRVGQPW/RTALALRSLYRARPSLRC
		١		- 		PPVELPWAPRRGHRLSPADDELYQRTRISLLQ
l						REAAQAMYIDSYNSRGFMINGNRVLGPCALL
	1					PHSVVQWNVGSHQDITEDSFSLFWLLEPRIEI
						······································

NO. of No. of hod DINO: peptide exide exide of peptide in uncloated in USSN per peptide in the per per per per per per per per per pe	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
mucleoide seq			l .		_		D=Asnartic Acid F=Glutamic Acid
Sequence							
Sequence							
914	seq-						
amino acid residue of peptide residue of peptide residue of peptide sequence T-Threenine, V=Valine, W=Typroplan, Y=Tyrone, X=Unknow, Y=Siop codon, P=possible nucleotide deletion, V=possi			1				O=Glutamine R=Arginine S=Serine
Periodic of Peptide Sequence Y=Tyrosine, X=UhAnown, *=Siop codon, Pepsisible nucleotide delicin, *=Possible nucleotide insertion Y=VorTiGRTERIO_SOVI_OAMRORGIAVEVQ DTPMACATPNIT_CHEGRYTGAALIPPPGGTS. TSLO_QAAQ TSLO_A							T=Threonine, V=Valine, W=Tryptophan
peptide		[ĺ		residue of		
		ļ			peptide	•	/=possible nucleotide deletion. \=possible
DITPHACATIFIC							
DITPHACATIFIC					· - · · · · · · · · · · · · · · · · · ·		VVVGTGDRTERLOSOVLOAMRORGIAVEVO
1940 A 4492 1 472 FFFETERSSVAQAGVQWRPLGSLQAPPGFT				1			
1940							
	590	1940	Α	4492	1	472	
SP1						l	
SP1			í		Í	1	RFSRDGLDLLT/S/GDPPTSASQSAGITGVSHR
1941 A 4495			ļ	1			
591 1941 A 4495 1444 1116 IAARTILAKTWNQLKRRTIMDSIKKTRIVITY 4 LKDNWVEDTIPQGAVPCTATAEGMKRLLFAL EWYADTERNEMSFAGGWVAPCTATAEGMKRLLFAL EWYADTERNEMSFAGGWVAPCTATAEGMKRLLFAL 592 1942 A 4496 2 919 RIRPLFSGRFTRPVCTMSDERRLPGSAVGWL 593 1943 A 4496 2 919 RIRPLFSGRFTRPVCTMSDERRLPGSAVGWL 593 1943 A 4506 2 193 FIFFEASSCSPQAGVGYDLGWLHAPPFYGSC 694 1944 A 4506 2 193 FIFFEASSCSVPQAGVGRPDLGWLHAPPFYGSC 695 1944 A 4507 1327 647 KMAGGVRPLRGLRALCRVLLFLSQFCILSGG 594 1944 A 4507 1327 647 KMAGGVRPLRGLRALCRVLLFLSQFCILSGG 595 1945 A 4512 533 264 FFFRMESS'SVARLECSGAISAPCNLHLIGSNN 596 1946 A 4513 3 1674 HASDHLYPNFLVNELLRQKQRFEERFKLD 597 1946 A 4513 3<							
	591	1941	Α	4495	1444	1116	
				1	Ì		
1942 A 4496 2 919 RTRPLFSGRFTPPVCTMSDERRIPGSAVGWL VCGGLSLLANAWGLSVGAKQKKWKPLEFL LCTLAATHMLNVAVPIATYSVQLRGRPFDF EWNEGLCKVPVSTFTLIATCTSVTSLSYHK MWMVCWPVNYRLNAKKQAGHTVMGTWM GSFILSALPA/GWHDTSERTPTHGCRFTVAEI GLGFGVCFLLLVGGSVAMGVICTALFOTL AVQVGRQAGHRAFTVPTIVVEDAQGKRRSSI DGSEPAKTSLQTTGLVTTIVPEDAQGKRSSI DGSEPAKTSLQTTGLVTTIVPEDAQGKRSSI DGSEPAKTSLQTTGLVTTIVPEDAGGKCVM FFFEAESCSVPQAGVQRPDLGWLHAPPPGGC HFPASASQVAGTTHARHHTQLIPAFLVENGL C C C C C C C C C	i .						
VCGGLSLLANAWGILSVGARQRKWRPLEFL							
LCTLAATHMINNAVPIATYSVOQLRRCRPDF	592	1942	Α	4496	2	919	RTRPLFSGRPTRPVCTMSDERRLPGSAVGWL
LCTLAATHMINNAVPIATYSVOQLRRCRPDF	1		1	1			VCGGLSLLANAWGILSVGAKQKKWKPLEFL
			İ				LCTLAATHMLNVAVPIATYSVVQLRRQRPDF
GSFILSALPAN VGWHDTSERFYTHGCRFIV AEI GLGFOVCFILLY GESVAMGVICTAIALFOTI. AVQVGRQADHRAFTVPITIVVEDAQGKRRSSI DGSEPAKTSI.QTTGL.VTITIVYEDAQGKRRSSI DGSEPAKTSI.QTTGL.VTITIVYEDAGGKRRSSI DGSEPAKTSI.QTTGL.VTITIVYEDCL.MGFPVL GFSLADTHL.SQLPYTWGDRDSGGACVM FFFEAESCSVPQAGVQRPDL.GWLHAPPPGSC HFPASASQVAGTTHARHITQLFVAFVENGL C HFPASASQVAGTTHARHITQLFVAFVENGL C GSTEIPPYVMKCPSNGLCSSLPADCIDCTTINFS CTYGKPVTFDCA VKPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDYECTINSTSCMTVSCPRQ RYPANCTVR.DHVHCLGNRTFPKML.YCNWT GGYKWVGJWLLRHHPRWGLGADRFYLLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRVAGNIGARRHITQQIFVLLVQMRVH YVGQDGLDLLNLMHHPPSFVL.GLQA HASPHLYPMFLVNGLKQQFFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAQLQLILMEFLK VARNKREQLEQIGUS.SVLEEDIKRVEMS GLYSPVSEDSTVPQFEAPSPSISIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLUS KATRYNSVRPLWTLSVASDLYNGSQYKSLV FEFDRDCDYFAIAGSUS.SVLEEDIKRVEMS DYEGTVILWDGFTGQRSKVYQEHEKKCWSV DFNLMDPKLLASSGDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHALAFGCADHCV HYYDLLRNTKQPIMVFKGHRKAVSYAKPVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIACGSSINSLYLYYKGLS KTLLTFKPDTVKSVLDKDRKEDDTTNEFVSAV CWRALPDGESNVLIAANSQGTIKVLELV CWRALPDGESNVLIAANSQGTIKVLELV CWRALPDGESNVLIAANSQGTIKVLELV CWRALPDGESNVLIAANSQGTIKVLELV CWRALPDGESNVLIAANSQGTIKVLELV CWRALPDGESNVLIAANSQGTIKVLELV SASQVAEITSVRHTTWLIFCTLQQQMGFHHVGE QAGLELLTSWDPDALASGGIGNSPHAWPP	1						EWNEGLCKVFVSTFYTLTLATCFSVTSLSYHR
							MWMVCWPVNYRLSNAKKQAGHTVMGIWM
							GSFILSALPAVGWHDTSERFYTHGCRFIVAEI
S93							GLGFGVCFLLLVGGSVAMGVICTAIALFOTL
DGSEPAKTSLQTTGLVTTVFTYDCLMGFPVL						ļ	AVQVGRQADHRAFTVPTIVVEDAOGKRRSSI
593 1943 A 4506 2 193 FFFEAESCSVPQAGVQRPDLGWLHAPPPGGC HFPASASQVAGTTHARHHTQLIF\AFLVENGL C C 594 1944 A 4507 1327 647 KMAGGVRPLRGLRALCRVLLFLSQFCILSGG ESTEIPPYVMKCPSAVRPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDVECTINSTSCMTVSCPRQ RYPANCTVRIDHVHCLGNRTFPKMLYCNWT GGYKWYQLWLRHHPRWGLGADRFYYLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRVAGNIGARHHTQQIFVLLVQMRVH YVQQDGLDLLNLMIHPPRSPKVLGQA 596 1946 A 4513 3 1674 HASDHLYPNFLVWELIILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEASHAAQLQILMEFLK VARRNREQLEQIQKELSVLEEDIKRVEEMS GLYSVESDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSQTKKQPWYNSTLASRKRATAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECULS KFITRYNSVRPLATLSYASDLYNGSQYKSLV FEFDDCDYFEAPSPSHSSIIDSTEYSQP PGFSGSQTKKQPWYNSTLASRKRATAHFE DLEQCYFSTRMSISDDSRTASQLDEFQECULS KFITRYNSVRPLATLSYASDLYNGSQYKSLV FEFDDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMITCNSKISCISWSSYHKINLLASS DVEGTVILWDGFTGQRSKVYQEHERCWSV DFNLMDPKLLAGSSDDSKSVKVLWSTNLDNSV ASIEAKANVCVKFSSSRYHLAFGCADHCV HYYDLRNIKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIACVSENSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMSAHCNLHLIGSSDPPTS ASQVAEITSVRHHTWLFCLUGOMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A <td< td=""><td></td><td></td><td>[</td><td></td><td></td><td></td><td>DGSEPAKTSLQTTGLVTTIVFIYDCLMGFPVL</td></td<>			[DGSEPAKTSLQTTGLVTTIVFIYDCLMGFPVL
593 1943 A 4506 2 193 FFFEAESCSVPQAGVQRPDLGWLHAPPPGGC HFPASASQVAGTTHARHHTQLIF\AFLVENGL C C 594 1944 A 4507 1327 647 KMAGGVRPLRGLRALCRVLLFLSQFCILSGG ESTEIPPYVMKCPSAVRPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDVECTINSTSCMTVSCPRQ RYPANCTVRIDHVHCLGNRTFPKMLYCNWT GGYKWYQLWLRHHPRWGLGADRFYYLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRVAGNIGARHHTQQIFVLLVQMRVH YVQQDGLDLLNLMIHPPRSPKVLGQA 596 1946 A 4513 3 1674 HASDHLYPNFLVWELIILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEASHAAQLQILMEFLK VARRNREQLEQIQKELSVLEEDIKRVEEMS GLYSVESDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSQTKKQPWYNSTLASRKRATAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECULS KFITRYNSVRPLATLSYASDLYNGSQYKSLV FEFDDCDYFEAPSPSHSSIIDSTEYSQP PGFSGSQTKKQPWYNSTLASRKRATAHFE DLEQCYFSTRMSISDDSRTASQLDEFQECULS KFITRYNSVRPLATLSYASDLYNGSQYKSLV FEFDDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMITCNSKISCISWSSYHKINLLASS DVEGTVILWDGFTGQRSKVYQEHERCWSV DFNLMDPKLLAGSSDDSKSVKVLWSTNLDNSV ASIEAKANVCVKFSSSRYHLAFGCADHCV HYYDLRNIKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIACVSENSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMSAHCNLHLIGSSDPPTS ASQVAEITSVRHHTWLFCLUGOMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A <td< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td>GPFSLADTHLSDLPYTWGDRDSGGACVM</td></td<>							GPFSLADTHLSDLPYTWGDRDSGGACVM
HFPASASQVAGTTHARHHTQLIFAFLVENGL	593	1943	Α	4506	2	193	FFFEAESCSVPQAGVQRPDLGWLHAPPP\GSC
1944 A 4507 1327 647 KMAGGVRPLRGLRALCRVLLFLSQFCILSGG ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNPS CTYGKPVFDCAVKPSYTCVDQDFKSQKNPII NMTCFEWQLPETDYECTNSTSCMTVSCPRQ RYPANCTVRDNYHCLGNRTFPKMLYCNWT GGYKWVYGLWLLRHIPRWGLGADRPIYLGP VAGTASGKLFSFGGLGIWTLLIDVLLIGVGYVG PADGSLYI 595							HFPASASQVAGTTHARHHTQLIF\AFLVENGL
ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ RYPANCTVR\DHVHCLGNRTFPKMLYCNWT GGYKWVYGLWLLRHIPRWGLGADRFYYLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLL/NLMIHPPRSPKVLGLQA 596 1946 A 4513 3 1674 HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSTNGHRWQIFQDWLGTDQDNLDLANV NLMELLVQKKKQLEAESHAAQLQLMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIDISTEYSQP PGFGGSSQTKKQPWYNSTLASRKKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KATIRYNSVRPLATILSYASDLYNGSQYKSLV FEFDRDCDYFALAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMYFKGHKAVSVAKFVSG EEIVSASTDSQLLUMVGKFPYCLRSFKGHIN EKNFVGLASNGDYLACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANSQGTIKKVLELV 597 1947 A 4518 536 824 RSLALSFGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCLLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 1 384 FPIEFVNIGGDFDAAAGVFRXCRLPGGAYFFSF			l				C
ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ RYPANCTVRUDHVHCLGNRTFPKMLYCNWT GGYKWVYGLWLLRHHPRWGLGADRF\YLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRV/AGNIGARHHTQGIFVLLVQMRVH YVGQDGLDLL/NLMIHPPRSPKVLGLQA 596 1946 A 4513 3 1674 HASDHLYPNFLVNELILKQKQRFEEKFFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIDSTEYSQP PGFGGSSQIFKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KFITRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSRYHLAFGCADHCV HYYDLRTKQPIMVFKGHKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANSQGTIKKVLELV 597 1947 A 4518 536 824 RSLALSFGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCLLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 1 384 FPIEFVNIGGDFDAAAGVFRXCRLPGGAYFFSF	594	1944	Α	4507	1327	647	KMAGGVRPLRGLRALCRVLLFLSQFCILSGG
NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ RYPANCTVRUDHVICLGNRTTPKMLYCNWT GGYKWVGLWLLRHHPRWGLGADRF\YLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRVIAGNIGARIHITQQIFVLLVQMRVH YVGQDGLDLL/NLMIHPPRSPKVLGLQA 596 1946 A 4513 3 1674 HASDHLYPNTLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKRQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSFVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQYKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KFITRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFALAGVTKKIKVYEYDTVIQDA VDIHYPPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLENTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYJACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF			ŀ				ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS
RYPANCTVR.DHVH.CLGNRTFPKML.YCNWT GGYKWVYGLWLLRHIPRWGLGADRFYLGP VAGTASGKL.FSFGGLGIWTLIDVLLIGVGYVG PADGSLY1							
GGYKWVYGLWLLRHIPRWGLGADRF\YLGP	- i		[l	•		NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ
VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI				1			RYPANCTVR\DHVHCLGNRTFPKMLYCNWT
PADGSLY							
595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHILLGSNN SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLL/NLMIHPPRSPKVLGLQA 596 1946 A 4513 3 1674 HASDHLYPNFLVNELILKQKQRFEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KF\TRYNSVRPLATLSYASDLYNGSQYKSLV FFFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHKAVSYAKFVSG EEIVSASTDSQLKUNVGKP\YCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHILGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							
SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLL/NLMIHPPRSPKVLGLQA 596 1946 A 4513 3 1674 HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KF\TRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHILLGSSDPPTS ASQVAEITSVRHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							
96 1946 A 4513 3 1674 HASDHLYPNFLVIELILKOKORFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KFATRYNSVRPLVATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPT'S ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	595	1945	Α	4512	533	264	
1946 A 4513 3 1674 HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KF1RYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLASPGLECSGMISAHCNLHLGSSDPPT'S ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							SPASASRV/AGNIGARHHTQQIFVLLVQMRVH
HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KFATRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLIGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPFGAYFFSF	50.5	40.5	L	<u> </u>	_		
NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECUS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNILLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 I 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	596	1946	A	4513	3	1674	
VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECUS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNILLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDREDDTINEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNIHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF				1			HSVSSTNGHRWQIFQDWLGTDQDNLDLANV
GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KF\text{Trynsvrplatlsyasdlyngsqykslv} FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							NLMLELLVQKKKQLEAESHAAQLQILMEFLK
PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							VARRNKREQLEQIQKELSVLEEDIKRVEEMS
DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							
KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF				1			PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE
FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCNLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF				1			DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS
VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSIALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCNLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							
DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							
DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCNLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF						1	VDIHYPENEMTCNSKISCISWSSYHKNLLASS
ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCNLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							DYEGTVILWDGFTGQRSKVYQEHEKRCWSV
HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							DFNLMDPKLLASGSDDAKVKLWSTNLDNSV
EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF						İ	ASIEAKANVCCVKFSPSSRYHLAFGCADHCV
EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTI\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	[HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG
S97 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP S98 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF			1				EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN
S97 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP S98 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							EKNFV\GLASNGDYIACGSENNSLYLYYKGLS
597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV
597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							CWRALPDGESNVLIAANS\QGTI\KVLELV
ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	597	1947	A	4518	536	824	RSLALSPGLECSGMISAHCNLHLLGSSDPPTS
598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							ASQVAEITSVRHHTWLIFCILGOMGFHHVGE
598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF					_		QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP
TLGKLPRKTLSVKLMKNRDEVOAMIYDDGSS	598	1948	Α	4524	1	384	FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF
	L i						TLGKLPRKTLSVKLMKNRDEVQAMIYDDGSS

SEQ ID	SEQ ID	Met	SEQ .	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ŀ			peptide	_	/=possible nucleotide deletion, \=possible
ł		ļ		sequence	l	nucleotide insertion
						RRREMQSQSVMLALRRGDAVWLLSHDHDG
						YGAYSNHGKYITFSGFLVYPDLAPAAPPGLG
						ASELL
599	1949	Α	4526	366	776	MGQPAPYAEGPIQGGDAGELCKCDFLVFTSP
		1				NPEAVCEAGTPAMFQTAWRQMESCSI/AQAG
1			İ			VQWRDPGSLHPPPLGFKRFSCLSLPSSWDYK
ļ		l	ļ	}		HAPPHPANFCIFSRDQVSPCWPGWSRSLDLVI
1			j			PPPWLPKVLGLQA
600	1950	Α	4529	776	334	FFFETESCYVAQAGVQWCDLCSLQAPPPG\SS
]			DPPASASRVAGTTGARHHTQLIFVFLVETGFH
1	1	l				\MLARDGLKLLTSSDPPASASQSSWDYRREPP
1					}	RLANFFVFLVETGSRYVAQAGVQWLFTGAIP
[Ĺ				LLISTGVLTCSVSDLGRFTPP
601	1951	Α	4533	1460	403	HEVQESIHFLESEFSRGISDNYTLALITYALSS
				1	1	VGSPKAKEALNMLTWRAEQEGGMQFWVSSE
ł]	1			SKLSDSWQPRSLDIEVAAYALLSHFLQFQTSE
1		}				GIPIMRWLSRQRNSLGGFASTQDTTVALKALS
1		1]			EFAALMNTERTNIQVTVTGPSSPSPVKFLIDT
1						HNRLLLQTAELADGTANGSV/SISANGFGFAI
1		1				CQLNVVYNVKASGSSRRRRSIQNQEAFDLDV
	Ì	1	1			AVKENKDDLNHVDLNVCTSFSGPGRSGMAL .
						MEVNLLSGFMVPSEAISLSETVKKVEYDHGK
	١.	1				LNLYLDSVNETQFCVNIPAVRNFKVSNTQDA
		İ				SVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSD
			L			VQRLPSL
602	1952	A	4540	1963	295	MRAPGRPALRPLPLPLLLLLLSSPWGRAVPC
1		1	1			VSGGLPKPANITFLSINMKNVLQWTPPEGLQG
1 .		i				VKVTYTVQYFIYGQKKWLNKSECRNINRTYC
				1	ļ	DLSAETSDYEHQYYAKVKAIWGTKCSKWAE
1					•	SGRFYPFLETQIGPPEVALTTDEKSISVVLTAP
i				Í		EKWKRNPEDLPVSMQQIYSNLKYNVSVLNT KSNRTWSOCVTNHTLVLTW\LEPNTLYCVHV
				1		ESFVPGPPRRAOPSEKQCARTLKDQSSEFKAK
	1				ļ	
	I				1	
	ļ					IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK
					•	IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INF\\TL
					·	IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG
						IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\
					•	IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI
	:					IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL
					-	IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE
					-	IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE
					-	IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY
603	1052		4543	3	600	IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN
603	.1953	A	4543	3	600	IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL
603	.1953	A	4543	3	600	IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI
603	.1953	A	4543	3	600	IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INF\\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVLEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL
603	.1953	A	4543	3	600	IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INF\\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVLEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE
603	.1953	A	4543	3	600	IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\
603	.1953	A	4543	3	600	IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INF\\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVUEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN
						IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV
603	1953	A	4543	3	938	IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN
						IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW
						IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISD WWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG
						IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS
						IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS QGQQTQSSAYGSSYTYPPSSLGGTVVDGQPG
						IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS QGQQTQSSAYGSSYTYPPSSLGGTVVDGQPG FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA
						IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INF\ITL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS QGQQTQSSAYGSSYTYPPSSLGGTVVDGQPG FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA VKTVGSVVSSVALTGVLSGNGGTNVNMPVS
						IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS QGQQTQSSAYGSSYTYPPSSLGGTVVDGQPG FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SPQAAPQPQQVAQPLPAQPPALAQPQYQSPQ
			-			QPPQ QVAQLEI AQITALAQIQIQIQ
605	1955	A .	4553	2	2304	ILLQEKRNCLLMQLEEATRLTSYLQSQLKSLC ASTLTVSSGSSRGSLASSRGSLASSRGSLSSVS FTDIYGLPQYEKPDAEGSQLLRFDLIPFDSLGR DAPFSEPPGPSGFHKQRRSLDTPQSLASLSSRS SLSSLSPPSSPLDTPFLPASRDSPLAQLADSCE GPGLGALDRLRAHASAMGDEDLPGMAALQP HGVPGDGEGPHERGPPPASAPVGGTVTLRED SAKRLERRARRISACLSDYSLASDSGVFEPLT KRNEDAEEPAYGDTASNGDPQIHVGLLRDSG SECLLVHVLQLKNPAGLAVKEDCKVHIRVYL PPLDSGTPNTYCSKALEFQVPLVFNEVFRIPV HSSALTLKSLQLYVCSVTPQLQEELLGIAQIN LADYDSLSEMQLRWHSVQVFTS\LNHQGRGR LGVQERAPPGTLHTPSPSPA/STDAVTVLLAR TTAQLQAVERELAEERAKLEYTEEEVLEMER KEEQAEAISERSWQADSVDSGCSNCTQTSPPY PEPCCMGIDSILGHPFAAQAGPYSPEKFQPSPL KVDKETNTEDLFLEEAASLVKERPSRRARGSP FVRSGTIVRSQTFSPGARSQYVCRLYRSDSDS STLPRKSPFVRNTLERRTLRYKQSCRSSLAEL MARTSLDLELDLQASRTRQRQLNEELCALRE LRQRLEDAQLRGQTDLPPWVLRDERLRGLLR EAERQTRQTKLDYRHEQAAEKMLKKASKEI YQLRGQSHKEPIQVQTFREKIAFFTRPRINIPPL PADDV
606	1956	A	4555	3429	776	PGSGPGPAPFLAPVAAPVGGISFHLQIGLSREP VLLLQDSSGDYSLAHVREMACSIVDQKFPEC GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL
607	1957	A	4563	1	4499	SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT TVRCMHLLLEAVPAVAPQTSILDLRFNRIREI QPGAFRRLRNLNTLLLNNNQIKRIPSGAFEDL ENLKYLYLYKNEIQSIDRQAFKGLASLEQLYL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion HFNQIETLDPDSFQHLPKLERLFLHNNRITHL VPGTFNHLESMKRLRLDSNTLHCDCEILWLA DLLKTYAESGNAQAAAICEYPRRIQGRSVATI TPEELNCERPRITSEPQDADVTSGNTVYFTCR AEGNPKPEIIWLRNNNELSMKTDSRLNLLDD GTLMIQNTQETDQGIYQCMAKNVAGEVKTQ EVTLRYFGSPARPTFVIQPQNTEVLVGESVTL ECSATGHPPPRISWTRGDRTPLPVDPRVNITPS GGLYIQNVVQGDSGEYACSATNNIDSVHATA FIIVQALPQFTVTPQDRVVIEGQTVDFQCEAK GNPPPVIAWTKGGSQLSVDRRHLVLSSGTLRI SGVALHDQGQYECQAVNIIGSQKVVAHLTVQ PRVTPVFASIPSDTTVEVGANVQLPCSSQGEP EPAITWNKDGVQVTESGKFHISPEGFLTINDV GPADAGRYECVARNTIGSASVSMVLSVNVPD VSRNGDPFVATSIVEAIATVDRAINSTRTHLF DSRPRSPNDLLALFRYPRDPYTVEQARAGEIF ERTLQLIQEHVQHGLMVDLNGTSYHYNDLVS PQYLNLIANLSGCTAHRRVNNCSDMCFHQKY RTHDGTCNNLQHPMWGASLTAFERLLKSVY ENGFNTPRGINPHRLYNGHALPMPRLVSTTLI GTETVTPDEQFTHMLMQWGQFLDHDLDSTV VALSQARFSDGQHCSNVCSNDPPCFSVMIPPN DSRARSGARCMFFVRSSPVCGSGMTSLLMNS VYPREQINQLTSYIDASNVYGSTEHEARSIRD LASHRGLLRQGIVQRSGKPLLPFATGPPTECM RDENESPIPCFLAGDHRANEQLGLTSMHTLW FREHNRIATELLKLNPHWDGDTTIYYETRKIVG AEIQHITYQHWLPKILGEVGMRTLGEYHGYD PGINAGIFNAFATAAFRFGHTLVNPLLLPGLD ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL RGLFGVAGKMRVPSQLLNTELTERLFSMAHT VALDLAANIQRGRDHGIPPYHDYNVCNLS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLR DGDRLWYENPGVFSPAQLTQIKQTSLARILCD
						ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL RGLFGVAGKMRVPSQLLNTELTERLFSMAHT VALDLAAINIQRGRDHGIPPYHDYRVYCNLS AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID
608	1958	A	4566	354	1135	FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAL IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKL GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD EFTIAATDDNHIFAWGNGGNGRLAMTPTERP HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILI VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGE GGPDAW
609	1959	A	4567	1	412	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH RARPRINLRNVIYSFAVTYCLNYISLAMSSTL KLSFHVLSGS
610	1960	A	4570	697	467	ECRGVISAH/CCTLCLPSSSDSASAF/RVARTT GTCDYAQLIFAFLVEMGFHHVGQDGLHLL/N LVIRPPRPPKVLGLQA

Sequence	SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide .location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
1961	eotide	seq-	ļ		location		
amino ecid of peptide residue of peptide sequence T-Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, Y=Stop codon, \(\) possible nucleotide deletion, \(\) possible nucleotide deletion, \(\) possible nucleotide deletion, \(\) possible nucleotide deletion, \(\) possible nucleotide deletion, \(\) possible nucleotide deletion, \(\) possible nucleotide insertion \(\) and \(\) deletion \(\) and \(\) deletion \(\) and	, -	uence	İ				M=Methionine, N=Asparagine, P=Proline,
	пепсе			914	_		
				ļ			
1961			ļ			Sequence	/=possible nucleotide deletion \=nossible
WNPNYPESPRIPAPELEKRUSGATTAGAALM	1		l	l		}	
LCAATAVLLSAQGGPVQSKSPRFASWDEMU LAGUALUALGAQGGPVQSKSPRFASWDEWL LAGUALUALGAQGGCANTIGAHPGSAERAGAR LSAGGSACQGTEGISTDLPLAPESRVDPVLAP LAGUALAQNSRIQQCHEKVAQORRLEKQHL LOTQLKAQNSRIQQCHEKVAQORRLEKQHL LOTQLKAQNSRIQQCHEKVAQORRLEKQHL LOTQLKAQNSRIQQCHEKVAQORRLEKQHL LOTQLKAQNSRIQQCHEKVAQVGQLGATTVPSG QSGLFEIQPQGSPPLVNCKMTSDGGVTYUGGE RIDGSVDPNPWEAYKAGGIDPHGEFVLGI REVHSITIGDRNSRLAVQLRDWDGMAELLJOS VHLGGEDTAVSLQLTAPVAGQLGATTVPSG LSVPFSTWDQDHDLRRDKNCAKSLGGWWF GTCSHSNLNGQVFRSIPQQRQKLKKGFWKT WKRYTPLQATTNLIQPMAAEAS SYSASSVAGITGTRHHRTRG SPFSASSVAGAGTGTRHHRTRG SPFSASSVAGAGTGTRHHRTRG SYSASSVAGTGTRHHRTRG SYSASSVAGTGTRHHRTRG SYSASSVAGTGTRHHRTRG SYSASSVAGTGTRHHRTRG SYSASSVAGGTGTRHHRTRG SYSASSVAGTGTRHHRTRG SYSASSVAGTGTRHHRTRG SYSASSVAGTGTRHHRTRG SYSASSVAGTGTRHHRTRG SYSASSVAGTGTRHHRTRG SYSASSVAGTGTRHHRTRG SYSASSVAGTGTRHATHGT SYSASSVAGTGTRHATHGT SYSASSVAGTGTRHATHGT SYSASSVAGTGTRHATHGT SYSASSVAGTGTRHATHGT SYSASSVAGTGTRHATHGT SYSASSVAGTGTHATHG	611	1961	A	4571	25	1396	ADPHTTVIRFFPAASATKRVLPPVLRVSSPRT
							LCAATAVLLSAQGGPVQSKSPRFASWDEMIN
LOTQLKAQNSHQQLFHKVAQQQRHLEKQH RIQHQSQFGLLBIKHLDHEVAKQQQRHLEKQH RIQHQSQFGLLBIKHLDHEVAKPARKRRIP EMAQPVPPAHNVSRLHRIPRDCQELFQVGSPE EMAQPVPPAHNVSRLHRIPRDCQELFQVGSP SEVEPSTWDQHDLRFDKDQATVVER RIDGSVDFNRPWEAYKAGFGDPHGEFWLGL RKVHSITGDRNSRLAVQLRDWDGNABLIGFS VHLGGEDTAYSIQLITAPVAQQLGATTVPPSG GICSHNLINGQYFRSPQQRGKLKKGFWKT WRGRYYPLQATTHLIQPMABAAS 612							
RIQHLOSOFGILDHIKHLOHEVÄKPÄRKRÄP EMAQPYDPÄHNYSILHALPÖLÖGLEPQÜGER GSGLFEIQPÖGSPPFLVNICKMTSDGGWYTVIQR RIPLOSOYDPINRYMEAYKAGPGDPHGEPWLGI, EKVHSITODRNSELAVQLRDWDGNAELLOFS VHUGGEDTAYSLQLTAPVAGQLGATTVPPSG LSVPFSTWDQDHDLRIDRIXGAKSLSGGWWP GTCSHSNLNGQYFRSIPQQRQKLKKGFWKT WRGNYPTQJATTMLIQPMAAEAAS GSPASASPVAGTTMLIQPMAAEAAS GSPASASPVAGGVQWRDLSSLQPPPGSR GSPASASPVAGGVQWRDLSSLQPPPGSR GSPASASPVAGGVQWRDLSSLQPPPGSR SSNSPASASQVAGRUPARHAGHIFVFLVEPRF HYGRAGLGFL/NLAICLPQHPKYLGLQACN LNIKPHPAHKYSIMGPNVHEMSYHIYI LSVPFSTWDQAGVAGPPARHAGHIFVFLVEPRF HYGRAGLGFL/NLAICLPQHPKYLGLQACN LNIKPHPAHKYSIMGPNVHEMWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHTHWQYRLJGS GGLFCAWVGTLLVAMATDHWQYRLJGS GGLFCAWVGTLLVAMATDHTHWQTLAMATDHTH	1						
EMAQPVDAHNVSRIHRIPRDCQEIFQVGER QSGLFEIPQGESPFLVNCKTHOGOWTVTOR RHDGSVDFNRPWEAYKAGFGDPHGEFWLGI EKVHSITGDRNSRLAVQLRDWDGNAELLQFS VHLGGEDTAYSLQLTAPVAGQLGATTVPPSG LSVPFSTWDQDHDLRRDKNCAKSLSGGWWF GTCSHSNINGGYFRSIPQQRGKLKKGFWKT WRGRYYPLQATTMLIQPMAEAAS LSVPFSTWDQDHDLRRDKNCAKSLSGGWWF GTCSHSNINGGYFRSIPQQRGKLKKGFWKT WRGRYYPLQATTMLIQPMAEAAS GSPASASPVAGITGTRHIRTRG GGLPCAWVCTILLVAMATDHWMQYRLSGS FAHQEL WRYCLGNKCYLQTDSIAYWAATTA FMILSALCAISIMGIMAPPVALITHGAVATTA FMILSALCAISIMGIMAPPVALITHGAVATTA FMILSALCAISIMGIMAPPVALITHGAVAGA GGLPAWVCTLAVAATATHA TRANSPSWVPALADEWNTLHQEVTT TRLPAGSQEPVKD GSPASASPVAGVAGCPGSGEBAVALVVHLEK ETGRLRQQVSSPVHRKHSPIGAAWEVADFQ PEQVETOPRAVSBEEPGSLISHGHOGLNIKK ERRPLPPNARPSPWVPALADEWNTLHQEVTT TRLPAGSQEPVKD GSPASASPVAGVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGASPVAGAASPVAGAASPVAGASPVAGASPVAGASPVAGASPVAG	1			ļ			
REDGSVDPNRPWEAYKAGFGDPHGEFWLÖL	İ						
VHLGGEDTAYSLQLTAPWGQLGATTVPSGC			}				
LSVPTSTWDQDHOLRDKNCAKSLGGWWF]]	İ			EKVHSITGDRNSRLAVQLRDWDGNAELLQFS
GTCSHSNINĞQYFRSIPQRQKILKKGFWKT	1						VHLGGEDTAYSLQLTAPVAGQLGATTVPPSG
			j			•	
1962	j			}			
GSPASASPVAGITGTR-HHRTRG	612	1962	A	4575	162	3	
1963							
HHVGRAGLGFL/NLAICLPO;PRKVLGLQACN LNIKPHPAHKYISMIQFNVHFMCMSVHIYI 614 1964 A 4589 727 299 PGSAQSAQRGRGRRARAGSATQITMYSFMG GGLFCAWVGTILLVVAMATDHWMQYRLSGS FAHQGLWRYCLGHKCYLQTDBIAYWNATRA FMILSALCAISGIIMGIMAF/GWVAVLMITFA GIFYMCAYRVHECRRLSTPR 615 1965 A 4590 2 4114 TIPPERIQAWAQKQCYQSGEBAVALVVHLEK ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ PEQVETQPRAVSREPGSLHISGHQEQLNRKR ERRIPKNARPSPWAJALADEWNILHQEVTT TRLPAGSQEPVKD 616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQFLPPGFKRFSCL SLPSSWEYRCVPPRLANFVFLVEMGFLHVQEV AGLELPTSGDPALASQSAGITGVTTVPSQPG 617 1967 B 4595 84 478 XRHGLREPLLERRCAAASSFQHSSSLGRELPY DPVDTEGFGEGGDMQERFLFPSYILDPEPQPT REKQLQELQQQGEEBRQRQQREEBRQQNL RARSREIPVGHDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYPMCSTLKKPLARRCFPAHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIGLVIAMA MSLSIILHLLAGMGWWMIMENSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILLSLEVILLLIFLERKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKEDDSPCPTTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALIGLQINAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFFLFSAFGRALRYH TGSLAFGALLATVGRIEGESGYH TGSLAFGAGLLATVGRIEGESGYH RALIGLQINAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFFLFSAFGRALRYH TGSLAFGALLATVGRIEGESGYH TGSLAFGAGLLATVGRIEGERICHATHRIRIVQVDT APPLNYYWVPILTVIVGSYLLAHGFFSVYGMC VDTLFLLGKLLIVGSVGLAFFFFTHRIRIVQDF APPLNYYWVPILTVIVGSYLLAHGFFSVYGMC	613	1963	A	4584	687	321	PLAQRRPFLWVTVKTNGHIWGSSTYPHFWGS
LNIKPHPAHKYISMIQFNVHFMCMSVHIYI	İ		j				SNS/PASASQVAGIPNARHQARIIFVFLVEPRF
614 1964 A 4589 727 299 PGSAQSAQRGRGRRRARAGSATQITMYSFMG GGLFCAWVGTILLVVAMATDHWMQYRLSGS FAHQGLWRYCLGNKCYLQTDSIAYWNATRA FMILSALCAISGIIMGIMAR/GWVAVLMTFFA GIFYMCAYRVHECRRLSTPR GIFYMCAYRVHECRRLSTPR TILPEKIQAWAQKQCPQSGEEAVALVVHLEK ETGRLRQQVSSPVHREKHSLGAAWEVADFQ PEQVETQYRAVSREEPGSLHSGHQEQLNRKR ERRPLPKNARPSPWVPALADEWNTLHQEVTT TRLFAGSQEPVKD 616 1966 A 4592 773 488 DFALVAQAGVQWINLGSPQPLPPGFKRFSCL SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTYPSGPG AGLELPTSGDPPALASQSAGITGVTTYPSGPG AGLELPTSGDPPALASQSAGITGVTTYPSGPG AGLELPTSGDPPALASQSAGITGVTTYPSGPG PVDVDTEGFGEGGDMQERFLFPEYILDPEPQPT REKQLQELQQQEEERQQQREERRQQNL RARSREHPVVGHPPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAIHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIGLVIAMA MSLLSIILLHLAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV VLHLRQTWLAFMILSLEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPTTAKT CNPETFPSSNESRQCPNARCGFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKFDDLPAPPLFSAFGRALRYH TGSLAFGALLAIVQIIRVILEYLDORLKAAEN KFAKCLMTCLKCCFWCLEKFIKFINRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAYLDKV TDFLFLIGKLLIVAGVSVGLAFFFFTHIRRIVQDT APPLNYYWPILTVIVGSYLLAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKNKKAAES			}				HHVGRAGLGFL/NLAICLPQHPKVLGLQACN
GGLFCAWVGTILLVVAMATDHWMQYRLSGS FAHQGLWRYCLGNKCYLQTDSIAYWNATRA FMILSALCAISGIIMGIMAP/GWVAVLMIFFA GIFYMCAYRVHECRRLSTPR 615 1965 A 4590 2 414 TILPEKIQAWAQKQCPQSGEAVALVVHLEK ETGRIRQQVSSPYHREKHSPLGAAWEVADFQ PEQVETQPRAVSREEPGSLHSGHQEQLNRKR ERRPLPKNARPSPWVPALADEWNTLHQEVTT TRLPAGSQEPVKD 616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ AGLELPTSGDPPAASQSAGITGOVTTVPSGPG 617 1967 B 4595 84 478 XRHGLREPLLERRCAAASSFQHSSSLGRELPY DPVDTEGFGEGGDMQERFLFFEYILDPEPQFT REKQLGLQQQQEEERRQRQQREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAIHAYKGVL MVGNETTYEDGHGSRKNTDLVEGAKKANG VLEARQLAMRFEDYTVSWYWIIGLVIAMA MSLLSIILLHLAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVLVDLGFQTDFRV VHHLRQTWLAFMILISILEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIPDDSPCFTAKT CNPETFPSNESSRQCPNARCQFAYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAAFPLFSAFGRALRYH TGSLAFGALLIALIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IATYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGLAFFFFTHIRRIVQDT APPLNYYWPPLTVIVGSVLIAHGFFSVYGMC VDTLFLCGEDLERNDGSAERPYFMSSTLKKL LNKTNKKAABES	614	1064	A	4590	727	200	
FAHQGLWRYCLGNKCYLQTDSIAYWNATRA FMILSALCAISGIIMGIMAF/GWVAVLMTFFA GIFYMCAYRYHECRRLSTIR 615 1965 A 4590 2 414 TILPEKIQAWAQKQCPQSGEEAVALVYHLEK ETGRLRQVSPVHREKHSPLGAAWEVADFQ PEQVETQPRAVSREEPGSLHSGHQEQLNRKR ERRPLPKNARPSPWVPALADEWNTLHQEVTT TRLPAGSQEPVKD 616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPPRLANFVPLVEMGFLHVGQ AGLELPTSGDPPALASQSAGTTGVTTVPSGPG 617 1967 B 4595 84 478 XRHGLREPLLERRCAAASSFQHSSSLGRELPY DPVDTEGFGEGGDMQERFLFPEYLDPEPQPT REKQLQELQQQGEEERQRQQREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFFAHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDVTVSWYWIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMEINSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVHILLLIFLKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNESRQCPANGRCOFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDLDLAFPLFSAFGRALRYH TGSLAFGALLALVQIIRVILEYLDQRILKAAEN KFAKCLMTCLKCCFWCLEKFIKFINRNAYIM IAYGTINFCTSARNAFFLLMRNIIRVAVLDKV TDFIFLLGKLLIVGSVGILAFFFTHRIRIVQDT APPLNYWVPILTVVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNNDGSAERPYFMSSTLKKL LNKTNKKAAES	014	1904	^	4369	121	299	GGI FCANAGTII I VAAMATDUNAGADI GGE
615 1965 A 4590 2 414 TILPEKIQAWAÇKQCPQSGEEAVALVVHLEK ETGRLRQOVSSPVHREKHSPLGAAWEVADFQ PEQVETQPRAVSREEPGSLHSGHGQLNRKR ERRPLPKNARPSPWVPALADEWNTLHQEVTT TRLPAGSQEPVKD 616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPPRLANFVELVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG 617 1967 B 4595 84 478 XRHGLREPLLERCAAASSFQHSSSLGRELPY DPVDTEGFGEGDMQERFLEPFYLDPPDPT REKQLQELQQQGEEERQRQQRREERQQNL RARSREHPVVOHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWHIGLVIAMA MSILSIILHHLAGMGWWMIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSLEVIIILLIFLRKRILI AIALIKEASRAVGYWCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSNESRQCPNARCQFAFYGGESGYH RALLGQUTNAFMFWLANFWLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALLALVQURVILEYLDQRILKAAEN KFAKCLMTCLKCCFWCLEKFIKFIRNAYIM IAIYGTNCTSKANAFFFHILMRNIRVAVLDKV TDFIFLLGKLLIVGSVGLAFFFFTHRIRIVQDT APPLNYWVPILTVVGSYLIAHGFFSVYGMC VDTLFFLEGELERNDGSAERPYFMSSTLKKL LNKTNKKAAES							
GIFYMCAYRVHECRRLSTPR 1965 A 4590 2 414 TILPEKIQAWAQKQCPQSGEAVALVVHLEK ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ PEQVETQPRAVSREEPGSLHSGHQEQLNRKR ERRPLFKNARPSPWVPALADEWNTLHQEVTT TRLPAGSQEPVKD 616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPPRALANFVFLVEMGFLHVQQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG 617 1967 B 4595 84 478 XRHGLREPLLERRCAAASSFQHSSSLGRELPY DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT REKQLQELQQQQEEERQQQNL RARSREIPVVGHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERYPYMCSTLKKPLARRCFPAHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIGLVIAMA MSLLSIILHLLAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDERV YLHLRQTWLAFMILSILEVIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIPDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWAKRYPDDLSPCPFTTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWAKRYPDDLSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWAKRYPDDLSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWAKRYPDDLSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWAKRYPDDLSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH TGSLAFGALILAIVQIIRVLEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFINRNAYIM IAIYGTNFCTSARNAFFILMRNIIRVAVLDKV TDFILFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES							
ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ PEQVETQPRAVSREEPGSLHSGHQEQLNRKR ERRPLPKNARPSPWVPALADEWNTLHQEVTT TRLPAGSQEPVKD 616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPJRLANFVFLVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG 617 1967 B 4595 84 478 XRHGLREPLLERRCAAASSFQHSSSLGRELPY DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT REKQLQELQQQEEEERQRQQRREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSSLSILHILLAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYYMASTAVFLSTSNEAVYKIFDDSSCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFINRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES							GIFYMCAYRVHECRRLSTPR
PEQVETQPRAVSREEPGSLHSGHQEQLNRKR ERRPLPKNARPSFWVPALADEWNTLHQEVTT TRIPAGSQEPVKD 616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPPRLANFVFLVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG AGLELPTSGDPPALASQSAGITGVTTVPSGPG AGLELPTSGDPPALASQSAGITGVTTVPSGPG DPVDTEGFGEGGDMQERFLFFEYILDPEPQPT REKQLQELQQQQEEERQRQQRREERQQNL RARSREHPVVGHPDPALPPSGVNCSGCAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLAGIMGWVMIIMENSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHRQTWLAFMILSILEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGQINAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALLIAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFINNAYIM IAIYGTINFCTSARNAFFILLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES	615	1965	A	4590	2	414	TILPEKIQAWAQKQCPQSGEEAVALVVHLEK
616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ AGLELPTSGPPPALASQSAGITGVTTVPSGPG AGLELPTSGPPALASQSAGITGVTTVPSGPG AGLELPTSGPPALASQSAGITGVTTVPSGPG PPVDTEGFGEGGDMQERFLFPEYILDPEPQPT REKQLQELQQQQEEERQRQQRREERQQNL RARSREHPVGHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIGLVIAMA MSLLSIILLHLAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHRQTWLAFMIILSILEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALGLQIFNAFMFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALLATVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM LAYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRVQDT APPLNYYWVPILTVIVGSYLLAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES			ŀ				
616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ AGELPTSGDPPALASQSAGTGVTVTVPSGPG GELPY DPVDTEGFGEGGMQERFLFPEYILDPEPQPT REKQLQELQQQQEEERQRQQRREERRQQNL RASREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHHAYKGVL MYGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMEISELGYRF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVIIILLHFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALGQIFNAFMFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFNNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLLAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES	1						
616 1966 A 4592 773 488 DFALVAQAGVQWHNLGSPQFLPPGFKRFSCL SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ AGLELPTSGDPPALASQSAGITGVTTVPSGPG 617 1967 B 4595 84 478 XRHGLREPLLERRCAAASSFQHSSSLGRELPY DPVDTEGFGEGGDMQERFLFPEYLLDPEPQPT REKQLQELQQQGEEERQRQQRREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMERSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIPDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALLAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWPILTVIVGSYLLAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKNKKAAES			J				
AGLELPTSGDPPALASQSAGITGVTTVPSGPG 1967 B 4595 84 478 XRHGLREPLLERRCAAASSFQHSSSLGRELPY DPVDTEGFGEGGDMQERFIFPYILDPEPQPT REKQLQELQQQQEEERQRQQREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMENSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIPDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCF WCLEKFIKFINRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES	616	1966	A	4592	773	488	DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL
617 1967 B 4595 84 478 XRHGLREPLLERRCAAASSFQHSSSLGRELPY DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT REKQLQELQQQQEEERQRQQRREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLLAGIMGWWMIIMENSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFFLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES			<u> </u>				
DPVDTEGFGEGGDMQERFLFFEYILDPEPQPT REKQLQELQQQQEEEERQRQQRREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAHHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILHHLAGIMGWVMIIMENSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTIFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES	(17	1007		4505	0.4		
REKQLQELQQQQEEERQRQQRREERRQQNL RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAIHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLLIGKLLVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES	617	1907	В	4393	84	478	
RARSREHPVVGHPDPALPPSGVNCSGCGAEL HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAIHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMENSELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES]						
HCQDAR* 618 1968 A 4596 2945 1188 ARSRNSARGVYGMCVDTLFLCFLEDLERNDG SAERPYFMCSTLKKPLARRCFPAIHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMEISELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFTHRIIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES			ļ				RARSREHPVVGHPDPALPPSGVNCSGCGAEL
SAERPYFMCSTLKKPLARRCFPAIHAYKGVL MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMEI\SELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES							
MVGNETTYEDGHGSRKNITDLVEGAKKANG VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMEI\SELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNIYMVI IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES	618	1968	Α	4596	2945	1188	
VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMEI\SELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES							
MSLLSIILHLLAGIMGWVMIIMEI\SELGYRIF HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES			}	ļ]
HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV YLHLRQTWLAFMIILSILEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES							
YLHLRQTWLAFMIILSILEVIIILLIFLRKRILI AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES							HCYMEYSRLRGEAGSDVSLVDLGFOTDFRV
AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRXPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLMRNIIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES					İ		YLHLRQTWLAFMILSILEVIIILLLIFLRKRILI
CNPETFPSSNESRQCPNARCQFAFYGGESGYH RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES							AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI
RALLGLQIFNAFMFFWLANFVLALGQVTLAG AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES							AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT
AFASYYWALRKPDDLPAFPLFSAFGRALRYH TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES							PALL GLOENAEMEENT AND ALCOUTE AC
TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES		İ	ļ	1			
KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES				[
IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES							KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM
APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES				İ			IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV
VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL LNKTNKKAAES							TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT
LNKTNKKAAES			ļ				APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC
Larar Las Lagar Las Last Last Last Last Last Last Last	619	1969	A	4601	2	357	RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			' ' '	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
İ		Ì	İ	peptide	sequence	/=possible nucleotide deletion, \=possible
		l	}	sequence		nucleotide insertion
— —				Sequence	 	GLAYGHFSYEFSNHRDVVVDLQGWVTGNGK
j			}			
1		ł				GLIYLTDPQIHSVDQKVFTTNFGKRGIFYFFN
(20	1070	<u> </u>	1606			NQHVECNEICHRLSLTRPSMEKPCKS
620	1970	Α	4606	1	2415	MERLWGLFQRAQQLSPRSSQTVYQRVEGPR
l i						KGHLEEEEEDGEEGAETLAHFCPMELRGPEP
1		ł			ł	LGSRPRQPNLIPWAAAGRRAAPYLVLTALLIF
1		<u>'</u>	1		}	TGAFLLGYVAFRGSCQACGDSVLVVSEDVN
Į,)			}	YEPDLDFHQGRLYWSDLQAMFLQFLGEGRL
1					[EDTIRQTSLRERVAGSAGMAALTQDIRAALS
1		l	}			RQKLDHVWTDTHYVGLQFPDPAHPNTLHWV
[ĺ	[1	DEAGKVGEQLPLEDPDVYCPYSAIGNVTGEL
		1				VYAHYGRPEDLQDLRARGVDPVGRLLLVRV
					1	GVISFAQKVTNAQDFGAQGVLIYPEPADFSQ
		l				DPPKPSLSSQQAVYGHVHLGTGDPYTPGFPSF
					}	NQTQFPPVASSGLPSIPAQPISADIASRLLRKL
						KGPVAPQEWQGSLLGSPYHLGPGPRLRLVVN
	l .				1	NHRTSTPINNIFGCIEGRSEPDHYVVIGAORDA
				i		WGPGAAKSAVGTAILLELVRTFSSMVSNGFR
						PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL
1				·	•	HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSL
1						IESVLKQVDSPNHSGQTLYEQVVFTN\PSWD\
1					(AEVIRPLPM\DSSAY\SFTAFVGVPAVEFSFME\
						DDQ\AYPFLHTKEDTYENLHKVLQGRLPAVA
						QAVAQLAGQLLIRLSHDRLLPLDFGRYGDVV
						LRHIGNLNEFSGDLKARGLTLQWVYSARGDY
						IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV
						EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL
						DHLRLLRSNSSGTPGATSSTGFQ\ESRFRRQL\
1 1						ALLITWDACKGAANALSGDVWNIDNNF
621	1971	Α	4610	793	334	ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRG\
021	17/1	Α.	4010	193	334	
						NTLVLKQQTFIESARSIGASDMTVLLRHILPGT
1 1						GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP
1 1		·				EWGAMLNEARADMVIAPHVAVFPALAIFLTV
			-15-			LAFNLLGDGLRDALDPKIKG
622	1972	A	4614	2	820	LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ
						CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF
]	RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK
						SCVILLGLLLLYDVFFVFITPFITKNGESIMVEL
1						AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV
				i		VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL
]						LIAYCRRFDVQTGSSYIYYVSV\TVAYAIGMIL
						TFVVLG\LMKKGQPALLYLVPCTLITA/CQFV
						AWETVREMKKFWERVTS
623	1973	A	4619	17	691	TLVSVVEFVRRADLTREDLAPSSVDSGQAGF
						GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT
						DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ
						IGHFLCLVILVYCAEYINEAAAMNWRLFSKY
1					İ	OYFDSRGMFISIVFSAPLLVNAMIIVVMWVW
						KTLNVMTDLKNAQERRKEKKRRKED*GAA
1 1						AAWSLRPSRPPSAAPSAAVCVAWASFOLTHG
	ļ					
624	1074		4622	164	660	LKNRCFI
624	1974	Α	4622	164	668	VSCYTALQSIMNQPESANDPEPLCAVCGQAH
]						SLEENHFYSYPEEVDDDLICHICLQALLDPLD
						TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL
				İ		QHCKKSSILVNKLLNKLLVTCPFREHCTQVL
1				1		QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED
<u> </u>						CLSPGVHHCSEV
625	1975	Α	4625	474	473	CFLSPSPLLPPLLLSSSSSPSFPLPPPPTLLPSTLP
L						PPLLIPSS*LSP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
626	1976	A	4629	249	3	KLKGNECFCYHCNVCIFLMIKK*GLFLC*IYFI LFFET*SHSFTRLECSGTISAHCSLQLQGSSNSP ASASQVAGIAGTHH
627	1977	A	4635	1	301	FFFFETKPFFAPQAGGQGPSRGSLNPLPTGLK QFSGLTLSRSGNNGPRPPPRVNFGILRGNGVP PGGAG*PRPPDLRGPPGLAPPQGGNNGGDPP ARAYL
628	1978	A	4648	1357	782	KLFSSQRLFGPHIQAINPSFLLLSFFPS*LLAMR TVGNNAFILVFLVYRIVLLLF*HV*PAYFQPSK NKTAKINCN*RPFLFLVCYLL*AELHIGIFIANF YDCIPNKLNEHLWPKLLQSLIFHVDFCGFLHK VFYICFTEFLLFLYFL*LFIIKVSCSII*CSTICVF SYKSFAVIIFFVDNTRFFSFGF
629	1979	A	4660	18	999	HHELHTLELLQNPKEVLTRSEIQDVNYSLEAV KVKTVCQIPLMKEMLKRFQVAVNLAEDTAH PKLVFSQEGRYVKNTASASSWPVFSSAWNYF AGWRNPQKTAFVERFQHLSCVLGKNVFTSG KHYWEVESRDSLEVAVGVCREDVMGITDRS KMSPDVGIWAIYWSAAGYWPLIGFPGTPTQQ EPALHRVGVYLDRGTGNVSFYSAVDGVHLH TFSCSSVSRLRPFFWLSPLASLVIPPVTDRK*G FSSPDQNSFPVVQLRDTHPWALFCPSCLYPG WSIFWVSLTVPFGICPLCASQEAVPWEVGLA NGDGTGNFPRRFWEIFL
630	1980	A	4669	2	358	FFFFFETESHSVAQAGMQWRNLGSLPAPPPGF TPFFCLSLLNGWDYRRPPHILANFFVLLVETG FHDVGQDGLDLLTS*STPSASQSAEITGVSHC TRLKKIRFAKGHVEFFFESHVE
631	1981	A	4674	953	614	TPIRGTDDEHEECTVQEYSAGKNTCLRPGAV AHTCNPCTLGGRGRWIT*GSGVQDQPGPTWQ NPVFLERRPRALHSSPGLTTQRILWAQGLWV GAGSTGCSRGPRGEGVFREG
632	1982	A	4678	34	314	RSTHASGMISPSFGFMGHLLRLEFEILPSTPNP *LPSYQGEAAGSSLISHLQTFSPDLKGVYCTFP ASGLAPVPTHWTVSELSRSPVATATFC
633	1983	A	4696	1	1365	RTLGMEGERRASQAPSSGLPAGANGESPGG GAPFPGSSGSSALLQAEVLDLDEDEDDLEVFS KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP DLKDLFITVDEPESHVTTIETFTTYRIITKTSRG EFDSSEFEVRRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERFNDDFIETRRKALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMINFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP
634	1984	A	4708	421	158	SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRRKTLQFPCYH
635	1985	A	4709	42	341	YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED
636	1986	A	4721	2	351	EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS RSTTVTVA*LMQKLNLSMNDAYYIVIMKMSS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met bod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion ISPNFNSMDQPLDFQRTLGLRSPCYNRVPAQK MYFTTPSNHNAYQVDSVQST
637	1987	A	4726	664	253	NTGLTCSIQRKCGETQLYRREENRLILLLQDH LKSESFQVLTLSPRLEFSGLISAHCNLRLPGSS DSSASSSRAAGITGVHHHAWLIFFFLVETGFL HAG*AGLELLTSGDPPASASRSAGITGVSHHA RPRETRFL
638	1988	A	4734	24	592	GGMDSRVSGTTSNGETKPVYPVMEKKEEDG TLERGHWNNKMEFVLSVAGEIIGLGNVWRFP YLCYKNGGGAFFIPYLVFLFTCGIPVFLLETAL GQYTSQGGVTAWRKICPIFEGIGYASQMIVIL LNVYYIIVLAWALFYLFSSFTIDLPWGGCYHE WNTEHCMEFQKTNGSLNGTSENATSPVIEFW
639	1989	A	4743	1040	699	QGLTLLPRMECSATITAHCSLELPGSIDLPTSA \$*VARTTGTHHHPWLILVLLL*TWGSYYVAQ AGLELLGSSNLPAAMVSQSAQIIGHDHCAWA TSNHVLYTQEGLRRGKEG
640	1990	A	4771	527 _	2	GRIDCPHPATVLAQPIFIDACSVLGAYQGAQN WIRRRPCLPSGCLKMNREIGPLQHSLCCPGWS QTPGLKAILLRQPPK*LGLQMESHSCPPAWSA MARSRLTATSASQVQAILLPQPPGTTDSCSPS PDHEQQPLSWVLPPPQKDMNPREQQVALGP QAAALPWAVWRNDCFPR
641	1991	Α .	4780	16	473	RPSSQCGGIPTGWKKGLAPELSSELSSPPLPAR LQLAASPYFSPSWAECPQPVPAGTHATWCLA RVWARMTPPGPAGIPSHPLPPPPPERSVPIPSP FPARDSGSRQGHSTDRYKHTDAPRDAHRRVP QRDTDTGVHTGSGTHTHAHTPPEK
642	1992	A	4798	1	487	GYSFRCDIVDYSRSPTALRMARTCWLYYFSK FIELLDTIFFVLRKKNSQVTFLHVFHHTIMPW TWWFGVKFAAGGLGTFHALLNTAVHVVMY SYYGLSALGPAYQKYLWWKKYLTSLQLVQF VIVAIHISQFFFMEDCKYQFPVFACIIMSYSFM FLLLFLH
643	1993	A	4799	2	391	LMAFIEMHISGSLVYLKIKTKIYSYFSMLNFLL QEIPLSEILRISSPRDFTNISQGSNPHCFEIITDT MVYFVGENNGDSSHNPVLAATGVGLDVAQS WEKAIRQALMPVTPQASVCTSPGQGKDHSK Q*ASVCTSPGQGKDHSKQ
644	1994	A	4800	488	101	AYPLFAVHPVHTECVAGVVGRAYLLCALFFL LSFLGYCKAFRESNKEGAHSSTFWVLLSIFLG AVAMLCKEQGITVLVRAATWLGPAFSVCPFP SYKDIWGWPCLCGVLHAYIPLLV
645	1995	A	4805	458	126	LLWTTVLCQTPARPQSTMIHLGHILFLLLLPV AAAQTTPGERSSLPAFYPGTSGSCSGCGSLSL PLLAGLVAADAVASLLIVGAVFLCARPRRSP AQEDGKVYINMPGRG
646	1996	A	4817	47	1033	LQGDTWHLSFLSHFSRLHGGVPGRGLLEGNL LQPQAPGHDMTSIPFPGDRLLQVDGVILCGLT HKQAVQCLKGPGQVARLVLERRVPRSTQQC PSANDSMGDERTAVSLVTALPGRPSSCVSVT DGPKF*SSN*KRIANGLGFSFVQMEKESCSHL KSDLVRIKRLFPGHPAEENGAIAAGDIILGRE WEGPRKASSSRCRGSWAMQLSVQAGPSFAS YYPAAVEVLHLLRGAPQEVTLLLCRPPFGAL PELEQEWQTPELSADKEFTRATCTDSCTSPIL GSRGQLGGTVPPQMQGKAWGLRPESSQKAIR EGTMGAKTERVERPUNDERVER
647	1997	J A	4854	1044	335	PRVRGDWPLEKKKSNSNIHPIFSWCGSTDSKD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		i		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			1			IVMPTYDLTDSVLETMGRVSLDMMSVQANT
	\	1		ĺ	ľ	GPPWESKNSTAVWRGRDSRKERLELVKLSRK
				ŀ		HPELIDAAFTNFFFFKHDENLYGPIVKHISFFD
			1	İ		FFKHKYQINIDGTVAAYRLPYLLVGDSVVLK
			1			QDSIYYEHFYNELQPWKHYIPVKSNLSDLLEK
		ľ	ľ	i	1	LKWAKDHDEEAKKIAKAGQEFARNNLMGD
(40)	1000	<u> </u>	4067	2020		DIFCYYFQTFPRNMPIYK
648	1998	Α	4867	2030	837	AGMLPAVGSADEEEDPAEEDCPELVPMETTQ
						SEEEEKSGLGAKIPVTIITGYLGAGKTTLLNYI
						LTEQHSKRVAVILNEFGEGSALEKSLAVSQG
	ł	1		}	1	GELYEEWLELRNGCLCCSVKDNGLRAIENLM
				1		QKKGKFDYILLETTGLADPGAVASMFWVDA ELGSDIYLDGIITIVDSKYGLKHLAEEKPDGLI
						NEATRQVALADAILINKTDLVPEEDVKKLRT
		ŀ				TIRSINGLGOILETORSRVDLSNVLDLHAFDSL
		1	Ì		;	SGISLOKKLOHVPGTOPHLDOSIVTITFDVPG
						NAKEEHLNMFIQNLLWEKNVRNKDNHCMEV
		}				IRLKGLVSIKDKSQQVIVQGVHELYDLEETPV
		ł	ļ			SWKDDTERTNRLVLLGRNLDKDILKQLFIAT
		l	}	}	}	VTETEKQWTTHFKEDQVCT
649	1999	A	4873	226	189	DGVSLLLPKLGVQWAQYWAHWQPPLPGFKR
			1			FSCLSLRSSWD*KCAPPHPAFVFLVEMGFHRV
		İ		i		GQAGLELRTSGDPPASASQSAGITGVSHLA*P
						TSMPLLPFQRLCVYI
650	2000	A	4874	2	437	FFFLRRSFAFVAQAGVQWCDLGSPQPLPPGF
		ļ		ļ		K*FSCLSLPSSWDYRHAPPPCPS*FLYF**RQG
		1				FTMLARLVLNS*PHDLPTSPSQSAEIKGVSHR
		i	1			CPASFYLFLKYYLEAKFCA*GECAPSAGVGA
651	2001	A	4898	1201	771	GYKRGHKSCLLINCVVQI
001	2001	A	4898	1701	771	DAWGPETRLARILNPDSFIEPRPGRLPELEATR
						PHMEPKASCPAAAPLMERKFHVLVGVTGSV AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY
•						SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL
						RRWADLLLVAPLDANTLGKVASGICDNLLTC
			1			VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ
		1		[VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA
	1	1				EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP
				ŀ		LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA
		L			l	LSLPRGQACSRLNWTEGPGLSFFQPGEAAA
652	2002	Α	4927	1	611	FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA
	1	l		}		SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL
	[ļ				PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR
	ļ	1		}		GQSPIPSRASSPSCSWAQVPGVALARCAGVC
						KPGDSWRVAACISGRCCSRGRRRGSGPRNPE
	1					QSFRGAWGPSFWGSWKSQRELSAGGAQAWP
		<u> </u>	L	<u> </u>		LLGSAGSGLRGEA
653	2003	A	4965	2	283	FFFFI*DGVSLCHPGWNAVARSWLTATSASR
	[VQAVSCFRLPSSWDYRHATMPG*FF*YF**R
C54	2004	 	40.00	 	125	WGFTHLAILVLNS*PQVICPPWPPKVLTLQA
654	2004	A	4968	3 .	437	RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD
	[1		IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW
		1		I		DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP
	l	İ		Ì	1	PPCLTHLAAASCVVVWCGRWKRDSAECQCD
655	2005	 	4983	201	207	HSCSAVSQQEDRCRSSSCS
دده ا	2005	A	4763	201	397	MNNNTTCIQPSMISSMALPHYILLCIVGVFGN
656	2006	A	4988	332	159	TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC
050	2000	1	4700	332	139	LVHKDMYREFFEEEAQASNKHVTRCLTSLVI REVHIKTMR*HFLPIRLEKNKNNIKD
657	2007	В	5008	129	465	MAGMKTASGDYIDSSWELRVFVGEEDPEAES
				120	1 702	I MAGNETAGOD I EGG WELK I TO GEED LEADS

<u> </u>	CEO IN	N 4-4	CEO	Predicted	Destination of	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		[amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1	1.	peptide	1	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
						VTLRVTGESHIGGVLLKIVEQINRKQDWSDH
		1				AIWWEQKRQWLLQTHWTLDKYGILADARLF
		<u></u>				FGPQHRPVILRLPNRRALRLX*
658	2008	A	5017	1	292	FFFFKETESHSVTQAGVQWHDLGSLQPPPPGF
		1	İ			KRFSCLSLLSSWDYRCAPPHPANFVFLVETGF
L		ļ				HHVAQAGLKLLTL*SANLGLSTSLPIPLFILLS
659	2009	Α	5018	17	338	RGHGGKSLTGGTPGNWGDGLLVSEDWSHLIF
ļ			}			T*NSLVSPVLGKWSPCLQGPGLSAVHTWPWL
j		1			}	MAACWAVHVKTHMRPGLAVLPRLVLNSWS
	0010	ļ	5000			*AIILLWPPKALGLQA
660	2010	Α	5028	2	310	SRVDDFVGERRGGCDECLCGHRGLRAVPLG
				ļ		HPGHLCLQPPGGPA*FLDYCRGCCPHPVPGST
1		1		Ì	1	AGSCPRQKKTTPGPTVLCVCSFWIYQRGEPH HRTGARWNH
661	2011	A	5050	752	431	RQSCSSTQAKVQWFHYGPLQSQPPGLKQSSQ
001	2011	^	3030	132	451	LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQAS
		1				LELLGSSHPPTSASQSARITGVSHRAWPLK*F
1		ļ			J	NLNQYQTLTMN
662	2012	Α	5054	48	103	ELNNGPFQMPLCNGGNLAVTGSWADRSPLH
002	2012	l ^•	3034	70	103	EAASQGRLLALRTLLSQGYNVNAVTLDHVTP
					Ì	LHEACLGDHVACARTLLEAGANVNAITIDGV
l		1			1	TPLFNACSQGSPSCAELLLEYGAQAQLESCLP
						SPTHEGASKGHHECLDILISWGIDVDQEIPHSG
		ł	1			TPLYVACMAQQFHCIWNLIYAGAGVRKGKY
		1		į		WDTPLPGAGHQSTQKLE*LFAMVEIWQ
663	2013	A	5066	951	580	VRNS*SFAHCASVYKHHYMDGQTPCLFVSSK
1						ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSCA
						GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF
		<u> </u>	<u> </u>			WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ
664	2014	Α	5071	550	1	LSFIEVLSMEQVNKTVVREFVVLGFSSLARLQ
1				İ		QLLFVIFLLLYLFTLGTNAIIISTIVLDRALHTP
1		1.	[1		MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK
		Ī]	TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR
·						YMAICNPLRYSVLMGHGVCMGLMAAAWAC
	0015	<u> </u>	5054	406	(00	GFTVSLVTTSLVFHLPFHSSNQHE
665	2015	A	5074	496	692	QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG
					1	PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD
		1		ļ		HDQNEGFHCREECRILGHSDRCWMPRNPMPI
	1	1			1	RSKSPEHVRNIIALSIEATAADVEAYDDCGPT
	1	ļ)	J	1	KRTFATFGKDVSDHPAEERPTLKGKRTVDVT
1		I				ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL
		1			1	PTKPSVSYEIVDPGITARRC
666	2016	A	5080	408	248	IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL
1	2010	1^	2030	100	1	VGPKAQLSGLQLQPCLYKRR
667	2017	A	5081	129	247	DLTNSHFFLFDFOKTGPPLGGPKAQFSSLOLO
***	~~``	1 '`	550.	1	1~	PCVY*RR
668	2018	A	5086	852	233	NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF
555	2010	''	7030	332		RALPTTFADIENLKYLLFTRDASQPFYLGHTV
		1				IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE
	l	1	1		1	TCADQSVIWKLSEDKQLAICLKYAGVHAENA
1	1	1	{	ĺ	i	EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG
1		1				CCSDMAITFNGLTPQKMEVMMYGLYRLRAF
1	1	1	1			GHYFNDTLVFLPPVGSEND
669	2019	A	5101	1	329	PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG
		1		l		ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP
		1				RGCQHEAAPCPRGPGSDGLHHASAACASLPP
		1		<u> </u>		SPILPVLLPELGPL
670	2020	A	5102	3	547	DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP
				·		the state of the s

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DEL*NMNGRVDYLVTEEEINLTRGPSGLGFNI VGGTDQQYVSNDSGIYVSRIKENGAAALDGR LQEGDKILSVNGQDLKNLLHQDAVDLFRNA GYAVSLRVQHRLQVQNGPIGHRGEGDPSGIPI
671	2021	Ā	5105	672	400	FMVLVPVFALTMVAAWAFMRYRQQL RDGREELCLQQEPTLPSRICSSAPLLYFLFICPF VLLLLLLISLLCLYWKARKLSTLRSNTRKEKA
672	2022	A	5148	72	314	LWVDLKEAGGVTTNRMED*EEDECN IIYFSYNIFLKITELLNDVERLKQALNGLSQLT YTSGNPTKRQSQLIDTLQHQVKSLEQQLAVS NQAHGALQEYVLAPCS
673	2023	Α	5152	210	335	REILCSRIGRLNIV*MSLFPNLTCRLNAIPIKIPA NHFVEVT
674		A	5153	3	2953	LTEDOPFDILQKSLQEANITEQTLAEEAYLDA SIGSSQGFAQAQLHPSSSASFTQASNVSNYSG QTLQPIGVTHVPVGASFASNTVGVQHGFMQH VGISVPSQHLSNSSQISGSGQIQLIGSFGNHPS MMTINNLDGSQIILKGSGQQAPSNVSGGLLV HRQTPNGNSLFGNSSSSPVAQPVTVPFNSTNF QTSLPVHNIIIQRGLAPNSNKVPINIQPKPIQM GQQNTYNVNNLGIQQHHVQQGISFASASSPQ GSVVGPHMSVNIVNQQNTTKPVTSQAVSSTG GSIVIHSPMGQPHAPQSQFLIPTSLSVSSNSVH HVQTINGQLLQTQPSQLISGQVASEHVMLNR NSSNMLRTNQPYTGPMLNNQNTAVHLVSGQ TFAASGSPVIANHASPQLVGGQMPLQQASPT VLHLSPGQSSVSQGRPGFATMPSVTSMSGPSR FPAVSSASTAHPSLGSAVQSGSSGSNFTGDQL TQPNRTPVPVSVSHRLPVSSSKSTSTFSNTPGT GTQQQFFCQAQKKCLNQTSPISAPKTTDGLR QAQIPGLLSTTLPGQDSGSKVISASLGTAQPQ QEKVVGSSPGHPAVQVESHSGGQKRPAAKQ LTKGAFILQQLQRDQAHTVTPDKSHFRSLSD AVQRLLSYHVCQGSMPTEEDLRKVDNEFETV ATQLLKRTQAMLNKYRCLLLEDAMRINPPAE MVMIDRMFNQEERASLSRDKRLALVDPEGFQ ADFCCSFKLDKAAHETQFGRSDQHGSKASSS LQPPAKAQGRDRAKTGVTEPMNHDQFHLVP NHIVVSAEGNISKKTECLGRALKFDKVGLVQ YQSTSEEKASRREPLKASQCSPGPEGHRKTSS RSDHGTESKLSSILADSHLEMTCNNSFQDKSL RNSPKNEVLHTDIMKGSGEPQPDLQLTKSLET TFKNILELKKAGRQPQSDPTVSGSVELDFPNF SPMASQENCLEKFIPDHSEGVVETDSILEAAV
675	2025	A	5154	599	1880	NSILEC LKKMEPFSCDTFVALPPATVDNRIIFGKNSDR LYDEVQEVVYFPAVVHDNLGERLKCTYIEID QVPETYAVVLSRPAWLWGAEMGANEHGVCI GNEAVWGREEVCDEEALLGMDLVRLGLERA DTAEKALNVIVDLLEKYGQGGNCTEGRMVF SYHNSFLIADRNEAWILETAGKYWAAEKVQE GVRNISNQLSITTKIAREHPDMRNYAKRGW WDGKKEFDFAAAYSYLDTAKMMTSSGRYCE GYKLLNKHKGNITFETMMEILRDKPSGINME GEFLTTASMVFILPQDSSLPCIHFFTGTPDPER SVFKPFIFVPHISQLLDTSSPTFELEDLVKKKS HFKPDRRHPLYQKHQQALEVVNNNEEKAKI MLDNMRKLEKELFREMESILQNKHLDVEKIV NLFPQCTKDEIQIYQSNLSVKVSS

						(A. Alesia C. C. et in
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	İ	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l	}	i	İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
l		ł	1	peptide	l	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
676	2026	Α	5155	2	306	FFFLRRSLALSPRPDCGLQWRNLGSLQAPPPG
1		1				FTPFSCLSLPSSWDYRRPPPRPANFLYF**RRG
						FTLLARMVSIS*PHDPPASASQSAGITGVSHRA
1	1	1			}	RPT
677	2027	Α	5167	97	740	FFHSVDLLALEQSKTFYKPDWFDIVESEVKCC
		1				KEAVCVIDMSSFTEFEITSTGDQALEVLQYLF
İ						SNDLDVPVGHIVHTGMLNEGGGYENDCSIAR
ļ]	LNKRSFFMISPTDQQVHCWAWLKKHMPKDS
	ļ			J	_	NLLLEDVTWKYTALNLIGPRAVDVLSELSYA
		ļ				PMTPDHFPSLFCKEMSVGYANGIRVMSMTHT
		1	}			GEPGFMLYIPIEYRWGFTMLSTLVSNS
670	2028	—	5183	1919	2018	PALCRLRDDMTVCVADFGLSKKIYSGDYYRQ
678	2028	A	2183	1219	2016	
				1	[GRIAKMPVKWIAIESLADRVYTSKSDVWAFG
Ĭ	ĺ	ĺ	1		1	VTMWEIATRGMTPYPGVQNHEMYDYLLHG
						HRLKQPEDCLDELCKI**SPQSP
679	2029	Α	5190	39	499	RESQVKHFKMRKIDLCLSSEGSEVILATSSDE
		ļ		Į.		KHPPENIIDGNPETFWTTTGMFPQEFIICFHKH
	ļ					VRIERLVIQSYFVQTLKIEKSTSKEPVDFEQWI
1	1		ł			EKDLVHTEGQLQNEEIVAHDGSATYLRFIIVS
						AFDHFASVHSVSAEGTVVSNLSS
680	2030	Α	5204	541	92	EILAVLKLACGDISLNALALMVATAVLTLAPL
				1		LLICLSYLFILSAILRVPSAAGRCKAFSTCSAH
1				ŀ		RTVVVVFYGTISFMYFKPKAKDPNVDKTVAL
			ì			FYGVVTPSLNPIIYSLRNAEVKAAVLTLLRGG
					1	LLSRKASHCYCCPLPLSAGIG
681	2031	A	5207	10	247	VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI
061	2031	^	3207	1.0	247	KNESPLPGTLAHTCNTSTLGGRGRWIT*GREF
į.				l		DTSMANMVKPCLYRK
(02	2032	A	5210	2	231	FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F
682	2032	Α	3210	4	231	
ł	l	l	ł	1		SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC
						WPGWSRTPDLS
683	2033	A	5218	85	402	CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF
ł		1		i		MIILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG
i						YARCYTFLIF*LFL*IPADSPTDQEPKTVMLSK
<u> </u>			<u> </u>			QSESAI
684	2034	A	5220	1	194	NLMKEMQNLNSENHKTWEEYKDTK*IMSYF
	1	1		I		YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL
		1	1	1		TDS
685	2035	A	5228	260	440	LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED
				Į.		QKKTPQVPINFTELSKCS*S*KIMSGERE
686	2036	A	5239	79	508	GGEAAARAAKLSSPRPHRVGRRERGVGGMS
330 .		1	1	1	1	AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR
					1	KHSRPIVTVWERELRKAKPNRKLTFLYLAND
Ī			}	1		VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD
			1	1		
	2022	-	5044	 	420	ESCKKHLGRVLSIWEERS
687	2037	A	5244	1	428	MAAVVAATALKGRGARNARVLRGILAGATA
1	ŀ					NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI
						PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR
		1				EVDKDRVKQMKARQNMRLSNTGEYESQRFR
		<u> </u>	<u> </u>	<u></u>		ASSQSAPSPDVGSGVQT
688	2038	A	5249	1	1407	LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP
1]	1	1			SGDSDLATALHRLSLRRQNYLSEKQFFAEEW
1	1	1	1			QRKIQVLADQKEGVSGCVTPTESLASLCTTQS
		1		i		EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY
1		1	1		1	HWQQLAQPNLGTILDPRPGVITKGFTQLPGD
1		1	,			AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS
1	l	1	1	ĺ		KPVTGIFLPPITSAGGPVTVATANPGKCLSCT
	I	I	1	[1	
l .		l .	1	i	1	1 NSTFTFTTCRILHPSDITOVTPSSGFPSLSCGSS
						NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS GSSSSNTAVNSPALAYRLSIGESITNRRDSTTT

SEQ ID	SEQ ID	Met	Lero	1 Designed	1 20-31-4 3	
NO: of	NO: of	hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1 1100	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ	1	peptide	1-1	/=possible nucleotide deletion, \=possible
	Į.			sequence		nucleotide insertion
			 	 		FSSTMSLAKLLQERGISAKVYHSPISENPLQPL
ł	1		1		1	PKSLAIPSTPPNSPSHSPCPSPLPFEPRVHLSEN
						FLASRPAETFLQEMYGLRPSRNPPDVGQLKM
						NLVDRLKRLGIARVVKNPGAQENGRCQEAEI
						GPQKPDSAVYLNSGSSLLGGLRRNQSLPVIM
			<u> </u>		l	GSFAAPVCTSSPKMGVLKED
689	2039	Α	5254	2	2621	LSLFGSRALGRSGARAMAKAKKVGARRKAS
]		ļ		GAPAGARGGPAKANSNPFEVKVNRQKFQILG
	1	į .				RKTRHDVGLPGVSRARALRKRTQTLLKEYKE
	1	ĺ		1	ł	RDKSNVFRDKRFGEYNSNMSPEEKMMKRFA
ļ	}	l	1	Ì		LEQQRHHEKKSIYNLNEDEELTHYGOSLADIE
				l		KHNDIVDSDSDAEDRGTLSGELTAAHFGGGG
				l		GLLHKKTQQEGEEREKPKSRKELIEELIAKSK
	1	1			!	QEKRERQAQREDALELTEKLDODWKEIOTLL
	ļ)]	SHKTPKSENRDKKEKPKPDAYDMMVRELGF
		1		ł		EMKAQPSNRMKTEAELAKEEQEHLRKLEAE
	1	1		1	[RLRRMLGKDEDENVKKPKHMSADDLNDGFV
	1	i	i			LDKDDRRLLSYKDGKMNVEEDVQEEQSKEA
	1	1		ļ		SDPESNEEEGDSSGGEDTEESDSPDSHLDLES
						NVESEEENEKPAKEQRQTPGKGLISGKERAG
		1				KATRDELPYTFAAPESYEELRSLLLGRSMEEQ
						LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE
						YVGDLATDDPPDLTVIDKLVVHLYHLCQMFP
						ESASDAIKFVLRDAMHEMEEMIETKGRAALP
	İ					GLDVLIYLKITGLLFPTSDFWHPVVTPALVCL
	}	ł				SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS
]	1			QRFIPELINFLLGILYIATPNKASQGSTLVHPFR
		l				ALGKNSELLVVSAREDVATWQQSSLSLRWA
						SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM
						YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ
		1				ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK
	İ	ļ	ĺ			REFKGAVREIRKDNQFLARMQLSEIMERDAE
	1					RKRKVKQLFNSLATQEGEWKALKRKKFKK
690	2040	A	5261	1	304	FFFFVFLVETGFHHVGQAGLELLTSGDPPTW
	1]			ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL
	1					FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT
			[[SFVK
691	2041	A	5270	3	158	NCHTTHCTANWVHLPGTPPGWKIDGPAAAL
			L			EVLSSFFFFFLKFSYKPQNIV
692	2042	A	5282	56	1268	GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV
	1			ĺ		ERVLTFLPAKALLRVACVCRLWRECVRRVLR
			1 1	ł		THRSVTWISAGLAEAGHLEGHCLVRVVAEEL
]]	ļ		ENVRILPHTVLYMADSETFISLEECRGHKRAR
]		ļ			KRTSMETALALEKLFPKQCQVLGIVTPGIVVT
] [] [ſ	PMGSGSNRPQEIEIGESGFALLFPOIEGIKIOPF
	1		1 l	}		HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV
	ļ		1 1	ļ		FGYNCCKVGASNYLQQVVSTFSDMNIILAGG
					}	QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI
			1			QSATVLLNEDVSDEKTAEAAMQRLKAANIPE
	}		1			HNTIGFMFACVGRGFQYYRAKGNVEADAFR
]		1 1	1	}	KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN
<u> </u>	لــــا			l		EVKDDDLFHSYTTIMALIHLGSSK
693	2043	A	5301	362	507	EEIKERFGPGLVIYWYGFIQELDCNRERGILLK
604	2011					ACFPTNIVTLCHSIA
694	2044	A	5310	1	204	RVLTAINHTLKENLRKFYKGKKDKPLDLRPK
				}	l	KTRAMRRRLNMHEENLKTKKQHRKERLYPL
	<u> </u>	!			[RKYAAKA
695	2045	Α	5315	125	1596	ETRSTAVKSEVQVCISLLLCLEDRTMPKKAKP

NO. of NO. of No. of wence order of the coation of	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
ueace Sequence 09,496 corresponding 1914 no first arinin acid residue of peptide residue of peptide squence 1914 no first arinin acid residue of peptide squence sque			hod				D=Aspartic Acid, E=Glutamic Acid,
Sequence							
uence 914 ng to first amino acid caresidue of peptide residue of peptide peptide peptide Falteronian, V-Valine, W-Pripophan, Y-Tyrosine, X-Unknown, *-Stop odoton, Peptide peptide Falteronian, V-Valine, W-Pripophan, Y-Tyrosine, X-Unknown, *-Stop odoton, Peptide peptide Falteronian, V-Tyrosine, X-Unknown, *-Stop odoton, Peptide peptide Falteronian, V-Tyrosine, X-Unknown, *-Stop odoton, Peptide peptide Falteronian, V-Tyrosine, X-Unknown, *-Stop odoton, Peptide peptide peptide peptide peptide Falteronian, V-Tyrosine, X-Unknown, *-Stop odoton, Peptide pept	1	•					
amino acid residue of sequence sequence sequence of peptide sequence of peptide sequence of peptide sequence sequence of peptide sequence of pepti		uciico		-			O=Glutamine, R=Arginine, S=Serine,
Poptide A-possible nucleotide deletion, p-possible nucleotide insertion Indeatodic in							T=Threonine, V=Valine, W=Tryptophan,
					residue of		Y=Tyrosine, X=Unknown, *=Stop codon,
1050KEEGPAPCKOMKLEAAGGPSALNFDSP							
SSLFESILSPIKTETFKEFWEGKPLLIQRIDPDA					sequence		
LATYYGSLFKLTDLKSLCSRGMYYGRDYOL							
CRCVNGKKKV1.NDGKAHFLQLRDEPDUS							
ATIOFHOPORPKDEL WRIGENECTYGSLVGS NVYITTAGSGGLPHYDDVEVELQLGEGK WRLYHPTYPTAGSGGLPHYDDVEVELQLGEGK WRLYHPTYPTAGSGGLPHYDDVEVELQLGEGK WRLYHPTYPTAGSGGLPHYDDVEVELQLGEGK WRLYHPTYPTAGSGGLPHYDDVEVELQLGEGK WRLYHPTYPTAGSGGLPHYDDATFAGLARITYTIST YQNNSWGDFLLDTISGLYDTAKEDVELTTC PROLLLQVESTTYATRALSGFLATTALDRIEG TKELLSSDWKKDFIMHRLPPYSAGDGAELSTP GGKLPRLDSVVRLQFKDHIVLTVLPQDQDGD ETQEKMYTYHSLKNSRETHIMMGNEEPTEFH GLRPFLSHLDALKQIWNSRAISYKDLKLITIDE EKSLVLSLWTECLIGVV LMKXYLEAAELGEISDHTKLLRISSSQGTIET SLQDIDSRLSPGGSLADAWAHQEGTHYKDRN VEKLQVILINCMTBITYYOPKKDKAGERRLAYN EEQHKPDQKLLYHATTAAMTPLSCVKK YRAFLNKSEWIRKMLHLRQLISLTINCOFT EESKYQEYTHELQETLPOKMTASSGIKHT MTPIYPSSNTLVEMTLGHTLAKKLKEEMEGVVKE LAENNHLLESGGSLTMGGLRNVFTASSGIKHT MTPIYPSSNTLVEMTLGHKKLKEEMEGVVKE LAENNHLLESGGSTLMGGLRNVFTASSGIKHT MTPIYPSSNTLVEMTLGHKKLKEEMEGVVKE LAENNHLLESGGSTLMGGLRNVFTASSGIKHT MTPIYPSSNTLVEMTLGHKKLKEEMEGVVKE LAENNHLLESGGSTLMGGLGRNVGTASGGET PPOPCKCSCLSLSSSWDYRHLPHLASICGED PPOPCKCSCCLSLSSSWDYRHLPHLASICGED PPOPCKCSCCLSLSSSWDYRHLPHLASICGED PPOPCKCSCCLSLSSSWDYRHLPHLASICGED PPOPCKCSCCLSLSSSWDYRHLPHLASICGED PPOPCKCSCCLSLSSWDYRHLPHLASICGED PPOPCKCSCCLSLSSWDYRHLPHLASICGED PPOPCKCSCCLSLSSWDYRHLPHLASICGED PPOPCKCSCCLSLSSWDYRHLPHLASICGED PPOPCKCSCCLSLSSWDYRHLPHLASICGED PPOPCKCSCCLSLSSWDYRHLPHLASICGED PPOPCKCSCCLSLSSWDYRHLPHLASICGED PPOPCKCSCCLSLSSWDYRHPHLASICGED PPOPCKCSCCLSLSSKAD PPOPCKSSCCLSC VSSWPCSCCTCGEN PPOPCKSSCC VSSWPCSCCTCCCSLSCEP PPOPCKSCSCLSCSCORGACATE AGAPENS PPOPKT VSCROLADICALS PPOPCKSCCLSC PPOPKT VSCROLADICALS PPOPCKSCCLSC PPOPKT VSCROLADICALS PPOPCKSCCLSC PPOPKT VSCROLADICALS PPOPCKSCCLSC PPOPKT VSCROLADICALS PPOPCKSCCLSC PPOPKT VSCROLADICALS PPOPCKSCCLSC PPOPKT VSCROLADICALS PPOPCKSCCLSC PPOPKT VSCROLADICALS PPOPCKSCCLSC PPOPKT VSCROLADICALS PPOPCKSCCLSC PPOPKT VSCROLADICALS PPOPCKSCCLSC PPOPKT VSCROLADICALS PPOPKT VSCROLADICALS PPOPCKSCCLSC PPOPKT VSCROLADICALS PPOPKT VSCROLADICALS PPOPKT VSCROLADICALS PPOPKT VSCROLA							
NYTITPAGSGGLPHYDDVEVFILQLEGEKH WRLYPHYPLARESYSVAEREIGRAFINGLESTHVITES WRLYPHYPLARESYSVAEREIGRAFINGLESTHVITES YQNNSWGDFLIDTISGLVPDTAGLARSTHVITES YQNNSWGDFLIDTISGLVPDTAGLARSTHVITES YQNNSWGDFLIDTISGLVPDTAGLARSTHVITES TYQNNSWGDFLIDTISGLVPDTAGLARSTHVITES TYQNNSWGDFLIDTISGLVPDTAGLARSTHVITES TRELLSDMKKDFIMHRLPPYSAGDAGLSTF GKRLPKLDSVVRLQFKDHIVLTVLPDQDQSD ETQEKMYTYTHSLKNSRETHMMGNEETEFT GLRPPLSHLDALKQWINSPAISVKDLKLTTDE EKESLVLSLWTECLIQVV EKESLVLSLWTECLIQVV EKESLVLSLWTECLIQVV EKESLVLSLWTECLIQVV EKESLVLSLWTECLIQVV EKESLVLSLWTECLIQVV EKESLVLSLWTECLIQVV YARTAMTHTDECVK YEARINSPEGRINGHTYOPKKDKAERRLANN EEQHKEPKQKLYVHATKAMTHTDECVK YEARINSPEGRINGHT THTYPSSNTLVEMTLOMKIKAERRLANN EEQHKEPKQKLYVHATKAMTHTDECVK YARTAMSEWINKHMLHR, ROLLSLTNOCPDI EEEVSK YQEYTHELQETH POKMFTASSGIKHT MTTYPSSNTLVEMTLOMKIKAERBEGVVK YARTAMSEWINKHMLHR, ROLLSLTNOCPDI EEEVSK YQEYTHELQETH POKMFTASSGIKHT MTTYPSSNTLVEMTLOMKIKAERBEGVVK YARTAMSEWINKHMLHR, ROLLSLTNOCPDI EEEVSK YQEYTHELQETH YOKULABOWN THTYPSSNTLVEMTLOMKIKAERBEGV YESSWPOWSRTPDFR SERVEN WELLDEN THTYPSSNTLVEMTLOMKIKAERBEGV YSSWPOWSRTPDFR YSSWPOWSRT							
KPGDLLYPPRGTHPQADTPAGLABSTHVTIST							
YQNNSWGPILDTISGLVPDTAKEDVELRTG PROLLLQVESTVATRLSGLSTRAPRLEG TKELLSSDMKXDFIMHRLPPYSAGDGAELSTP GRKJPRLDSVVRLQFKDHIVLTVLPDQDQSD ETQEKMVYIYHSLKNSREHHMMONEETEHH GRFPLSHLDALKQIWNSPAISVKDLKLTTDE EKESLVLSLWTECLQVV]]						WRLYHPTVPLAREYSVEAEERIGRPVHEFML
PRQLLLQVESTTVATRRLSGEARTLADREG TEELLSSDMKKDPHMRLPPYSAGDGAELSTP GGKLPRLDSVVRLOFKDHIVLTVLPODQDSD FTQEKMYVYHSIKNSRETHMMGNEEFTEPH GRFPLSHLDALKQIWNSPAISVKDLKLTTDE GGKLPRLDSLVARUFCKHIVVTVLPODQDSD FTQEKMYVYHSIKNSRETHMMGNEEFTEPH GRFPLSHLDALKQIWNSPAISVKDLKLTTDE EKESLVLSLWTECLIQVV							
TKELLSSDMKKDFIMHRLPPYSAGDAELSTP							
GGKLPRLDSVVRLQPKDHIVLTVLPQDQSD ETQEKMYLYHSLKNSRIADKMGNEEFIEFH GLRFPLSHLDALKQIWNSPAISVKDLKLTTDE EKBSLVLSLWTECLIQVV							
ETQEKMYYIYHSLKNSRETHMMGNEETEFH GIRFPLSHILDALKQIWNSYKOLIKLITDE EKESLVLSLWTECLIQVV							
GLRPPLSHLDALKQIWNSPAISVKDLKLITDE							
696							
SLQDIDSRLSPGGSLADAWAHGEGTHPKDRN VEKLQVLINCMTEILYQPKDKDKAERRIA YN EEQHKFDKQKLYYHATKAMTHFTDECVKK YEAFLNKSEEWIRKMLHLRQLUSLTINQCFDI EEEVSKYQEYTNELQFOKMTASSGIKHT MTPTYPSSNTLVEMTLGMKKLKEEMEGVVKE LARNNHLBEGGSLTHDKGLRNVDCL 697 2047 A 5320 244 478 LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG VSPSWPGWSRTPDFR 698 2048 A 5324 266 714 LPIRKSLRSVRSGPFTSQSFITRNLDGTASGSG CALKYTVTGSLPRINVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETVQERKQKQRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALINLPRNFSVIDKODKD 699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA KGEALINLPRNFSVIDKODKD 699 2050 A 5344 3 614 PTAEEMSSLTPESSFELAKRSWFGNFISLDKEE VGRUNGKTTIVCPFFIKTGM FEGCTTGCPSLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSFELAKRSWFGNFISLDKEE VGRUNGKTTIVCPFFIKTGM FEGCTTGCPSLPLIEPKYAVEKIVEAILQEKM YLYMPKLLYFTTLISGPSRFIKDKEE VGRUNGKTTIVCPFFIKTGM FEGCTTGCPSLPLIEPKYAVEKIVEAILQEKM YLYMPKLLYFTTLISGPSRFIKDKEE VGRUNGKTTIVCPFFIKTGM FEGCTTGCPSLPLIEPKYAVEKIVEAILQEKM YLYMPKLLYFTLISGPSFRFIKN VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSPRRGPFICDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMGEDRDSSC STVGGVQTOSKOLLEPLSLPESSFGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV LADVKLVADPQRYILVYMRKAFTEQPTIDFGSV RINSTAFFEQENYPTLLCDVLPEDRLIREELQ KQRIREILEQQQQERNDTINTHGVCMFCCNEEF LGNRSVLINIMMAREHAPNIGLPDNIVNCEFL CTLQKLDNICALYFUNYLEELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLLVNFIRRQPCCCYGGCTVNFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFP TYENDTLLWTLSDESSDLTAQEQNENVPIISE DTSKLYVALKQSSLNQLLL							EKESLVLSLWTECLIQVV
VEKLQVLINCMTETYYOFKKDKAERRIAYN BEQHKEPOKQLLYHATKAMTHFTDECVKK YEAFINKSEEWIRKMLHLRKQLLSLTNQCFDI EEEVSKYQEYTNELOETLPQKMFTASSGIKHT MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE LAENNHILESGGSLTMDGGLRNVDCL 697 2047 A 5320 244 478 LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP PPGFKQFSCLSLPSSWDYRHLPHLANFSREG VSPSWPGWSRTPDFR 698 2048 A 5324 266 714 LPIRKSLRSVRSGFPTSQSPTTRNLDGTASGSC LAKTVTGSLFRNVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD 699 2049 A 5334 699 277 RPHGHLVCISSAGLGGVNGLADYCASKFAA PGFABSVPVETFVQKQKGIKTTVCPFFIKTGM FEGCTTGCPSLLPILEPKYSVEVKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPLAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKVPRQVDISSSEG PESPFRDGSGGGGGTSVTTTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFGRVMASKTQGAVARMQEDRDGSC STVGGVGYGDSNCLLEPLSLPESPGGTTTLE GSPSVPCIFCEBHPPVAEQDKLLKHMIEHKIV LADVSLVADFQRVILWRRFFTEQPTDFCSV RINSTAPFEQENYFLLCDVLPEDRILRELQ KQRLREILEQQQCERNDTNFHGVCMFCNEEP LGNRSVILNHMAREHAFNIGLPDNTVCNEEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKNPKNRSYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYYHMEDAHAFDLLKIKSELG LNFYQQVKLVNFRRQVHQCRCYGCTGVTKFKS KADLRTHMEETKHTSLLPDSESDLTAQEQNENVPIISE DTISKLYALKQSSILNQLLL	696	2046	Α	5318	1476	742	
### BEQHKEDKQKLYYHATKAMTHFTDECVKK YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI EEEVSKYQEYTNELQETLPQKMFTASSGIKHT MTPIYPSSNTLVEMTLGMKKLKEEMEGVVVKE LAENNHILESGGIKHDGGLRNYDCL 697 2047 A 5320 244 478 LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG VSPSWPGWSRTPDFR 698 2048 A 5324 266 714 LPIRKSLRSVRSGFPTSQSPTTRNLDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD 699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVYVETFVQKQKGIKTTIVCPFPIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEALLQEKM YLYMPKLLYPMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSPSLSHSIVLS QTSFRAEVKASGGFSVFQKPVRFQVDISSSEG PEPSPRDGSGGGGIYSVTFTLISGPSRFFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPFPGRPDELSSSPRGPFKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEHFPVACQKLLKHMIEHKIV LADVKLVADFQRYLWRKFFTEQPTTDFCSV RINSTAPFEGQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEFL CTIQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKIPKNREFYDRFYVNNYLELGKSWEE VQLEDDRELLDHOEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDSSEDLTAQEQNENVPHISE DTISKLYALKQSSILNQLLL							
VEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI EEVSKYQEYTNELQETLPQKMFTASSGIKHT MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE LAENNHILESGGSLTMCGGLRNVDCL LDYNFLFEMTFEGLVSQAGVQWHDLGSLQPP PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG VSPSWPGWSRTPDFR VSPSWPGWSRTPDFR LPIRKSLRSVRSGFPTSQSPTTRNLDGTASGSC LAKTVTGSLFRNVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETVQERKQLDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD KKIEALLNLPRNPSVIDKQDKD RPHGHLVCISSSAGLSQVNGLADYCASKFAA FGFAESVFVETTVQKQKGKITTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMREKSSPLLKTGLLIADYLGI LHAMDGFADQKK LHAMDGFADQKK TIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMREKSSPLKTGNIFISLDKEE QIFLVLKDKPLSSKADIVHAFLSPSLSHSVLS QTSFAEYXASGGPSVFQKPVKFQVDISSSEG PEPSPRRDGSGGGGIYSVFTILISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPKDKK LLATNGTPL LATNGTPL SPSVPCTFCEHEPYAQDKLLKHMIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV RINSTAPFEGENYFLLCDVLPEDRILREELQ KQRLREILEQQCAPACHQENLLKHMIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV RINSTAPFEGENYFLLCDVLPEDRILREELQ KQRLREILEQQCAPACHCHKMIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV RINSTAPFEGENYFLLCDVLPEDRILREELQ KQRLREILEQQCAPACHCHKMIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV RINSTAPFEGENYFLLCDVLPEDRILREELQ KQRLREILEQQCAPACHCHKMIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV RINSTAPFEGENYFLLCDVLPEDRILREELQ KQRLREILEQQCAPCHCNETKS KACHRKNPKNREYDRFYVNNYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KACHRKNPKNREYDRFYVNNYLELGKSWEE VQLEDDRELLDHQDAUSCHVFIKS KACHRKNPKNREYDRFYVNNYLELGKSWEE VQLEDDRELLDHQDAUSCHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KACHRKNPKNREYDRFYVNNYLEIGKSWEE VQLEDDRELLDHQALBCHORDVPIISE DTSKLYALKQSSILNQLLL DTSKLYALKQSSILNQLLL DTSKLYALKQSSILNQLLL DTSKLYALKQSSILNQLLL DTSKLYALKQSSILNQLLL DTSKLYALKQSSILNQLLL DTSKLYALKQSSILNQLLL DTSKLYALKQSSILNQLLL DTSKLYALKQSSILNQLLL DTSKLYALKQSSILNQLL DTSKLYALKQSSILNQLL DTSKLYALKQSSILN			Į ,			}	
BEEVSKYQEYTNELQETLPQKMFTASSGIKHT MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE LAENNHILESGGSLTMDGGLRNVDCL							
MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE							
697 2047 A 5320 244 478							
698 2048 A 5324 266 714 LPRKSLRSYRSGFPTSQSPITRNLDGTASGSC LAKTYTGSLFRINVGLRGLVAGGIIGALLGTP VGGLIMAFQKYSGFTTSQSPITRNLDGTASGSC LAKTYTGSLFRINVGLRGLVAGGIIGALLGTP VGGLIMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALINLPRNPSVIDKQDKD 699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFILKSFLPLKTGLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSSLSHSVLS. QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSPRGPFKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGGSKDCILEPLSLPESSPGGTTTLE GSPSVPCIFCEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILLYWKRRTEQPITDFCSV IRINSTAPFEGOENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGRRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLCYYCEKTFRQKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSD WEEHPASAVCL FCCKQAETTEKLYVHMEDAHFPDLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFF TYENDTLLWTLSDSSEDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL			1			j	
VSPSWPGWSRTPDFR	697	2047	A	5320	244	478	
698 2048 A 5324 266 714 LPIRKSLRSVRSGPFTSQSPITRNLDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLMAFQKYSGETVQERGQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD 699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEALLQEKM YLYMPKLLYPMMFLKSFLPLKTGLILADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGFSVFQKPVRFQVDISSSEG PEPSPRDGSGGGGIYSVTTILISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPFGRPDPELSSSPRRGPFKDK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESSFGGTTTLE GSPSVPCIFCEEHPPVAEQDKLLKHMIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEQENYFLLCDVLPEDRILRELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGRNSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFF TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL]		j		
LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGILMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD 699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVPQKPVRFQVDISSSEG PEPSPRROSGGGGIYSVTTFLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV LADVALVADFQRYILYWRKRFTEQPITDFCSV RINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQGERNDTNFHGVCMFCNEEF LGRISVILNIFMAREHAFNIGLPDNIVNCNEEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDKTIWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	100		ļ				
VGGLLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD	698	2048	Α	5324	266	714	LPIRKSLKSVKSGFF1SQSP11KNLDG1ASGSC
KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD 699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTITVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFAEAYKASGGPSVFQKPVFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGPPDELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHPVAEQDKLLKHMIEHFIKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRNSTAPFEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEP LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILLPDRKTWDQLEYYFP TYENDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		1					
KKIEALLNLPRNPSVIDKQDKD 699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGISVYFTLLSGPSRFFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTITLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV LADVKLVADFQRYILYWKRFTEQPITDFCSV IRNSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFROKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFP TYENDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		ļ					, , ,
FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRDGSGGGGTYSVTTTLISGPSRFFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFF TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	ì	,	1			}	
FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEFSPRRDGSGGGIYSVTIFTLISGPSRFFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKFFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFF TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	699	2049	Α	5334	699	277	
YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI			l				
LHAMDGFADQKK							FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM
700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENVFLLCDVLPEDRILREELQ KQRLEILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV LADVKLVADFQRYILYWRKFFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	700	2050	-	52//	1	614	
QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTIQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	/00	2030	^	2244	١	"	
PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLEIGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		1					QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG
AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGRRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV
TOTO TOTAL TOTAL TOTAL TOTAL TOTAL TOTAL TOTAL TOTAL TOTAL TO THE TOTAL							
701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	1		1		l .	[
STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		005:	<u> </u>	5246		1202	
GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	701	2051	l A	3346	ا ا	1383	
IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL					{		
IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	1	1			1	1	
KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	1		ļ		1	1	IRINSTAPFEEQENYFLLCDVLPEDRILREELQ
CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		1			1	1	KQRLREILEQQQQERNDTNFHGVCMFCNEEF
KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL			1	[
VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL			1			1	
FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL					1		
LNFYQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL					1		
KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL					1		
TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL					1		
DTSKLYALKQSSILNQLLL			}	1			
702 2052 A 5356 2502 1540 MAAATRGCRPWGSLLGLLGLVSAAAAAWD					<u> </u>	1	DTSKLYALKQSSILNQLLL
	702	2052	Α	5356	2502	1540	MAAATRGCRPWGSLLGLLGLVSAAAAAWD

nucleotide sequence uence uence uence in nucleotide cotide sequence uence uence in nucleotide ucleotide uence in nucleotide uc	TI-THOUGHE,
sequence 09/496 correspondi to last amino M=Methionine, N=Asparagin quence 914 ng to first acid residue Q=Glutamine, R=Arginine, S=	
uence 914 ng to first acid residue Q=Glutamine, R=Arginine, S	e, P=Proline,
	=Serine,
amino acid of peptide T=Threonine, V=Valine, W=7 residue of sequence Y=Tyrosine, X=Unknown, *=	
residue of sequence Y=Tyrosine, X=Unknown, *= peptide /=possible nucleotide deletion	
sequence nucleotide insertion	, —possioic
LASLRCTLGAFCECDFRPD	DLPGLECDLAOHL
AGQHLAKALVVKALKAF	
HGWTGTGKSYVSSLLAHY	
HFSPVLHFPHPSHIERYKK	
CGRSLFLFDEMDKMPPGL	MEVLRPFLGSSWV
VYGTNYRKAIFIFISNTGGI RRDREEILLQELEPVISRAV	
MEERLLDAVVPFLPLQRH	
GLEPRDEVVQAVLDSTTFI	
TVASRIAFFL	
703 2053 A 5380 278 657 LFLQKLRMKTEEEARTHT	
EERLEFWMEKYDKDTEM	KQNELNALKATKA
SDLAHLQDLAKMIREYEQ	
KVKQDLLELKSVIKLQAW KM	WRGIMIRREIGGF
704 2054 A 5381 1 1003 FRGRAVKMAAVVEVEVG	GGA AGERET DEV
DMSDLSPEEQWRVEHARM	
AEMVLILIATLVVAQLLLV	
MVTLFQMWVVPLYFTVK	
SAVTAFVTFRATRKPLVQ	ITPRLVYKWFLLIY
KISYATGIVGYMAVMFTL	
MDFGISLLFYGLYYGVLEF	
STIGFYSESGMPTKHLSDS' SEEGIIENTYRLSCNHVFHI	
QTCPYCKEKVDLKRMFSN	
LDWLRYLVAWQPVIIGVV	QGINYILGLE
705 2055 A 5396 3 675 IYDRDPLQLATRAGQPLDI	
GNKRPLSALYRLESKEPFL	
RDDFYNRLFDYHGRVPPPI	
VTTTRRGKGVFSMKGGSR DELQTIKKELTQIKTKIDSV	
AEAEAQKKLLEESLVLIQE	
AEGGPDADGEEMTDGIEE	AFDEDGGHELFLO
K	ì
706 2056 A 5410 2 98 GRVGLNLEGRGCSEPKWR	HCTPTWATEQDSI
S S	
707 2057 A 5415 6 287 PFKLTPSFLSHAFSSQERI	
VRGVFVLEEFGNYTILLLG EEGLGAGRKRTSVEKSGG.	
708 2058 A 5423 3 291 SSSNPLGSPSTLWKLCSFV	
TPTLRAITLTVRVCGFIPEV	
GCTIFKTVTLTARSTASLLI	
709 2059 A 5424 679 347 RIRHEEKRGSRGRGRRTSE	EDTPKKKKHKGG
SEFTDTILSVHPSDVLDMP	
VSYGEMIGCDNPDCPIEWN GKWFCPRCVQEKRKKK	HFACVDLTTKPK
710 2060 A 5442 1073 559 QESLKKKIQPKLSLTLSSSV	/SRGNVSTDDDUSS
GSLTPPVTPPITPSSSFRSST	
EESDSDESWTTESAISSEAI	LSSMCMNGGEEK
PFACPVPGCKKRYKNVNG	IKYHAKNGHRTQI
RVRKPFKCRCGKSYKTAQ	GLRHHTINFHPPV
SAEIIRKMQQ	
711 2061 A 5449 1 319 GDSLCVPQYNKYREERVII	FLKMASGHAFQP
DLVKRIRDAIRMGLSARHV NGKKVEVAVKQIIAGKAV	FOGGA FORDERS
LYRDIPELQGF	PAGGWLSINLEITD
712 2062 A 5499 91 749 RPTPGHGDFWMQPLTKDA	GMSLSSVTLASAL
QVRGEALSEEEIWSLLFLA	AEQLLEDLRNDSS
DYVVCPWSALLSAAGSLS	FQGRVSHIEAAPF

No. of N			T				
mucleotide muc	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Control Cont		1	hod				
1949 1949			ļ				
1914 ng ng first amino acid residue of peptide pep	eotide	seq-	i		location		
1	seq-	uence		09/496	correspondi		
anino acid residue of peptide residue of peptide residue of peptide sequence T-Thremine, V-Valine, W-Typoine, V-Stop codon, /-possible nucleotide deletion, V-possible nucleotide nucleotide nucleotide nucleotide nucleotide nucleotide nucleotide nucleo	uence]	}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
residue of peptide peptide pepti						of peptide	
peptide		1	İ		1		
						1 504-5	
KAPELL,GGSEDEGPDASGMHYYSLGMTLY WSAGPHYPHQPLQLC,CEPLHSILLTMCEDQPH RRCTLQSVLEACR/HEKEVSYYPAPAGLHIR RLVGL/UGTISEVSREPCYSSSCWAKK RRCTLQSVLEACR/HEKEVSYYPAPAGLHIR RLVGL/UGTISEVSREPCYSSSCWAKK RRCTLQSVLEACR/UGTISEVSREPCYSSSCWAKK WEELIL/SRLDPHILHTPMYFELAH.SFLDLSFT TSSPQLLYNLNGCOXINSIMGEAQGLFELGL GGVECLLJA/WAYDRCVAICKPLHYMAVDM PRLCRGL/SYTWGCOX/ANSLAMSPYTURLPR CGHHEVDHFLCEMPALIRMACISTV CGHHEVDHFLCEMPALIRMACISTV RLTSSVSCALDEAAAALTRMRAESTANAGQS RRCTLSSVSCALDEAAAALTRMRAESTANAGQS RRCTLSSVSCALDEAAAALTRMRAESTANAGQS RRCTLSSVSCALDEAAAALTRMRAESTANAGQS RRCTLSSVSCALDEAAAALTRMRAESTANAGQS RRCTLSSVSCALDEAAAALTRMRAESTANAGQS RRCTLSSVSCALDEAAAALTRMRAESTANAGQS RRCTLSSVSCALDEAAAALTRMRAESTANAGQS RRCTLSSVSCALDEAAAALTRMRAESTANAGQS RRCTLSSVSCALDEAAAALTRMRAESTANAGQS RRCTLSSVSCALDEAAAALTRMRAESTANAGQS RRCTLSSVSCALDEAAAALTRMRAESTANAGQS RRCTLSSVSCALDEAAAAALTRMRAESTANAGQS RRCTLSSVSCALSGASGARSKLSKSKRKARGLQR RRCTLSSVSCALSGASGARSKLSKSKRKARGLQR RRCTLSSVSCALSGASGARSHCSKLSKSKRKARGLQR RRCTLSSVSCALSGASGASGARSLEGPTEPEAPTQ LACVSTRURLGERQFRPTEREAPTQ LACVSTRURLGERQFRPTEREAPTQ LACVSTRURLGERQFRPTEREAPTQ LACVSTRURLGERQFRPTEREAPTQ LACVSTRURLGERQFRPTEREAPTQ LACVSTRURLGERQFRPTEREAPTQ LACVSTRURLGERQFRPTEREAPTQ LACVSTRURLGERQFRPTEREAPTQ LACVSTRURLGERQFRPTEREAPTQ LACVSTRURLGERGASTANATS RACVSTRURLGERGASTA		1	1				
WSAGEHVPPHQPLQLCPLHSILLTIMCEDOPH RRCTLQSVLLACRYHEKEVSVYPAGALHIR RLVGLVLGTISEVSREPCFSSSSCWSCVARK RVELLVSRUPRILLTIMEKEVSVYPAGALHIR RLVGLVLGTISEVSREPCFSSSSCWSCVARK VEELLVSRLDPHLTPMYFFLAHLSPLDLSFT TSSPQLLYNLNGCDKTISYMGCAQLFLFLGL GGVECLLAVMAYDRAVALKCRLHYMVIMN PRICRCQLVSVTWGCQVANSLAMSPVTLRLPR GGHEVDHFLCEMPALIRMACISTV GHHEVDHFLCEMPALIRMACISTV GHHEVDHFLCEMPALIRMACISTV RLTSSVSCALDEAAAALTRMRAESTANAGQS NK STORE NK RTSSVSCALDEAAAALTRMRAESTANAGQS NK STORE NK RTSSVSCALDEAAAALTRMRAESTANAGQS NK STORE		 	 		sequence		
RRCTLQSVLEACRYHEKEVSVYPAPAGLHIR RIVGL_VGTISEVSREPCFSSSSCWCVAIK RIVGL_VGTISEVSREPCFSSSSCWCVAIK RIVGL_VGTISEVSREPCFSSSSCWCVAIK RIVGL_VGTISEVSREPCFSSSSCWCVAIK RIVGL_VGTISEVSREPCFSSSSCWCVAIK GGVECLL_JA_WAYDRCVAICKPL_HYMVIDAN PLCRG_USVTW_VGCVANSL_AMSPVTL_PLR GGVECLL_JA_WAYDRCVAICKPL_HYMVIDAN PLCRG_USVTW_VGCVANSL_AMSPVTL_PLR GGHEV_DHFLCEMPALIRMACISTV GHEV_DHFLCEMPALIRMACISTV ARRYWCENNIDIGKLSTADGKAFADPVTL_PLR GHEV_DHFLCEMPALIRMACISTV ARRYWCENNIDIGKLSTADGKAFADPVTL_PLR GHEV_DHFLCEMPALIRMACISTV ARRYWCENNIDIGKLSTADGKAFADPVTL_PLR ARRYWCENNIDIGKLSTADGKAFADPVTL_PLR ARRYWCENNIDIGKLSTADGKAFADPVTL_PLR ARRYWCENNIDIGKLSTADGKAFADPVTL_PLR ARRYWCENNIDIGKLSTADGKAFADPVTL_PLR ARRYWCENNIDIGKLSTADGKAFADPVTL_PLR ARRYWCENNIDIGKLSTADGKAFADAGS DK		1	į]	J		KAPELLQGQSEDEQPDASQMHVYSLGM1LY
RIVGL/LITISEVSREPCESSSSCWSCVAIL 2063 A 5506 22 478 VEELLU/SRLDPH-HTMYFFIL-HIS/FLDLSFT 7581PQL_YNINGCDK/TISYMGCAQLO_FLF_LGL_GGVSCLLLAVMAYDRCVAICE/HYMVIMO FREE			1				WSAGFHVPPHQPLQLCEPLHSILLIMCEDQPH
713			-	ļ	}	1	
TSSIPQLLYNINGCDKTISYMGCAIQLFLFLGL GGVECLLLAVMAYDRCVAICKPIWMYMM PRICEGLILAVMAYDRCVAICKPIWMYMM PRICEGLILAVMAYDRCVAICKPIWMYMM PRICEGLILAVMAYDRCVAICKPIWMYMM PRICEGLIVSVTWGCGVANSLAMSPYTLRLPR CGHIEVDHELCEMPALIRMACISTY	·						
	713	2063	A	5506	22	478	VEELILVSRLDPHLHTPMYFFLAHLSFLDLSFT
		1	l	1	ł	Í	TSSIPQLLYNLNGCDKTISYMGCAIQLFLFLGL
PRI.CRGL/SVTWGCGYANSLAMSPYTLRPR		1	ļ				
CGHHEVDHFLCEMPALIRMACISTY			1				
RLTSSVSCALDEAAAALTRMRAESTANAGQS	714	2064		5514	25	220	
DK	/14	2004	A	3314	25	220	1
			ļ	1	ļ	l	
GRTALPHHSGGSSGYESLRRDSEATGSASSAP	·		L				
DSMSESGAASFGARTRSILSSPKKATGLQRR RILPAPL BOTTAL GRKPSLPGWVDLPPPLAG SLKEPFEIKVYEIDDVERLQRPRPTPREAPTQG LACVSTRIRLAERRQQRIREVQAKHKHLCEE LAETQGRIMLEPQRWILEQFEVDPELEPESAE YLAALERATAALEQCVNICKAHVMMYTCFD ISVAASAAPGPQEVDV	715	2065	A	5526	3	810	
DSMSESGAASFGARTRSILSSPKKATGLQRR RILPAPL BOTTAL GRKPSLPGWVDLPPPLAG SLKEPFEIKVYEIDDVERLQRPRPTPREAPTQG LACVSTRIRLAERRQQRIREVQAKHKHLCEE LAETQGRIMLEPQRWILEQFEVDPELEPESAE YLAALERATAALEQCVNICKAHVMMYTCFD ISVAASAAPGPQEVDV			1				GRTALFHHSGGSSGYESLRRDSEATGSASSAP
RLIPAPL PDITAL GRKPSL PGQWVDL PPP LAG SLKEPFEK VYEID VERL QR PRPTPREAPT QG LACVSTRL RLAERR QGRLREV QAKHKHL CEE LAET QGRLMLEPGR WLEQFE V PELEPES AE VLAALER AT AALEQ CVNLCK AH VMM VTCFD ISVAASAAIPGP QE VDV		1		1	1		
SLKEPFEIKVYEIDDVERLQRPRPTPREAPTOG LACVSTRLRAERRQQRIREVQAKHKHLCEE LAETQGRLMLEPGRVLEQFEVDPELEPESAE YLAALERATAALEQCVNLCKAHVMMVTCFD ISVASAAAPIPGQEVDF.			1	1	[ĺ	
LACVSTR.RLAERRQQR.REVQAKHKHLCEE			1	1			
LAETQORLMLEPORWLEGFEVDPELEPESAE				1		1	
1						İ	1 APTOCRI MI EDGOMI POPEMBRI EDECAR
		J]	j		,	
716]		1		
		<u> </u>					
KLHPSSI CLALTLLSSVQGLQSISGLRLTDTF	716	2066	A	5529	458	790	
LKRTYEYDDIAQVCV			1		· ·		
LKRTYEYDDIAQVCV		i .		1			KLHPSSSI CLALTLLSSVQGLQSISGLRLTDTF
717		İ		1			
RSYHDRKSK VDLDRLNDDAKRYSCTPRNYS VNIREELKLANVVFFPRCLLVQRCGGNCGCG TVNWRSCTCNSGKTVKKYHEVLOFEPGHIKR RGRAKTMALVDIQLDHHERCDCICSSRPPR 718 2068 A 5586 311 88 AVLKNMAPMTALGLDLHILNILFLSAGEDF TSVVSEIMMYILLVFLTULLIEMIYCYRKVS KAEEAAQENA 719 2069 A 5598 1 330 KNCANEAVVQKILDRVLSRYDVRLRPNFGSM LATNSTRGINEDELMAHGQEKDSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FFLAFGLFNIVAAERC 720 2070 A 5628 798 148 LPPAQIPEAWILLANVVVVILUPLKDRLIDP LLIRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHHNETVSQQIGEVLVNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLGGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFILLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELFTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFFRLTQVTDYJLLIFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLISTIDW GLLQPAPAGAVLWGAVRVEKNQELCHLISTIDW GLLQPAPAGAVLWGAVRVEKNQELCHLISTIDW GLLQPAPAGAVLWGAVRVEKNQELCHLISTIDW GLLQPAPAGACVWACPPGTYQTESWRCVTA ERCASLHSVPGRACTWGCSCQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRACTTGCHQGSCLAQCPSGFT	717	2067	A	5531	3	460	
VNIREELKLANVVFFPRCLLVQRCGGNCGCG TVNWRSCTCNSKTVKKYHEVLQFEPGHIKR RGRAKTMALVDIQLDHERCDCICSSRPPR RGRAKTMALVDIQLDHERCDCICSSRPPR TSVVSEIMMYTLLVFLTLWLLIEMIYCYRKVS KAEEAAQENA TSVVSEIMMYTLLVFLTLWLLIEMIYCYRKVS KAEEAAQENA TSVVSEIMMYTLLVFLTLWLLIEMIYCYRKVS KAEEAAQENA TSVSEIMMYTLLVFLTLWLLIEMIYCYRKVS KAEEAAQENA LATNSTRGLNEDELMAHGQEKDSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC TSVSEIMMYTLLVFLTLWLIEMIYCYRKVS KAEEAAQENA LATNSTRGLNEDELMAHGQEKDSSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC TSVSEIMMYTLLVFLTLWLIEMIYCYRKVS LATNSTRGLNEDELMAHGQEKDSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL PPLAFGLFNIVAAERC TSVSEIMMYTLLVFLTLWLIEMIYCYNDKWSRFL PPLAFGLFNIVAAERC TSVSEIMMYTLLVFLTLWLIEMIYCYNDKWSRFL PSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL PSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL PSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL LATNSTRGLNEDELMAHGQEKDASSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL PSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL LATNSTRGLNEDELMAHGQEKDASSSESEDSC PSPGCSFTEGFSFDLNPDYVPKVDKWSRFL LATNSTRGLNEDELMAHGQEKDASSSESEDSC PSPGCSFTEGFSFDLNPDYVPKVDKWSRFL LATNSTRGLNEDELMAHGQEKDASSSESEDSC PSPGCSFTEGFSFDLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSSLAGEGAVLNACH WQIPCYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVSGVSGVLALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYEASQGPASHSRFSRDRG TSVSEIMMYTLVFLUMICTURUS LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATLHAMACTARGECHATE LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPDYVPKVDKWSRFL LATNSTRGLNPOKLAGECHAVCHANA RDNTGKRYDPTVVSNKLAGECADVCVCPCH HGMACTARGECHTECLGCSQPEDRACV ACRHLYFQACHUMACAVPCENGUELHLISTIDW GLOPPADATIONATOR ACRHLYFQACHTONATOR ACRHLYFOATOR ACRHLYFO	,,,	2007	••	555.	١	100	
TVNWRSCTCNSGKTVKKYHEVLQFEPGHIKR RGRAKTMALVDIQLDHHERCDCICSSRPPR 718 2068 A 5586 311 88 AVLKIMAPMTALGLLDLHILNLILFLSAGEDF TSVVSEIMMYHLLVFLTLWLLIEMIYCYRKVS KAEEAAQENA 719 2069 A 5598 1 330 KNCANEAVVQKILDRVLSRYDVRLRPNFGSM LATNSTRGLNEDELMAHGQEKDSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC 720 2070 A 5628 798 148 LPPAQIPEAWLLLANVVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHINETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFILLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELFTVLGANVNAA KLHETALHHAAKVKNVDLIEMLEFGGNIYA RDNRGKRPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGGSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGHQGSCLAQCPSGFT		140		-	ł		
718 2068 A 5586 311 88 AVLKNMAPMTALGLLDLHILNLILFLSAGEDF TSVVSSEIMMYILLVFLTLWLLIEMIYCYRKVS KAEEAAQENA 719 2069 A 5598 1 330 KNCANEAVVQKILDRVLSRYDVRLRPNFGSM LATNSTRGLNEDELMAHQQEKDSSSESEDSC PPSPGCSTIEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC 720 2070 A 5628 798 148 LPPAQIPEAWLLLANVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHINETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFILAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRI.FLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGHQGSCLAQCPSGFT			1		i	1	
T18 2068 A 5586 311 88		1	1	1	ł	1	
TSVVSEIMMYILLVFLTLWLLIEMIYCYRKVS KAEEAAQENA 719 2069 A 5598 1 330 KNCANEAVVQKILDRVLSRYDVRLRPNFGSM LATNSTRGLNEDELMAHGQEKDSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC 720 2070 A 5628 798 148 LPPAQIPEAWLLLANVVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSIPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLI LSQLCFVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT			ļ	 _			
Total	718	2068	Α	5586	311	88	,
Total		1	1				
LATNSTRGLNEDELMAHGQEKDSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC 720 2070 A 5628 798 148 LPPAQIPEAWLLLANVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHINETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT		1			}		
PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC 720 2070 A 5628 798 148 LPPAQIPEAWLLLANVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHINETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFILAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	719	2069	A	5598	1	330	KNCANEAVVQKILDRVLSRYDVRLRPNFGSM
PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC 720 2070 A 5628 798 148 LPPAQIPEAWLLLANVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHINETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFILAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT		1	1	1	}		LATNSTRGLNEDELMAHGOEKDSSSESEDSC
FPLAFGLFNIVAAERC 720 2070 A 5628 798 148 LPPAQIPEAWLLLANVVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYPQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT		l		ı			
Total]		1	ł		
LLLRCKLLPSALQKMALGMFFGPTSVIVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	720	2070	Ι	5620	709	140	
LEMERLHYIHINETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	120	2070	^	3028	'70	140	
WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT		[ļ	1	ſ	[
AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT			1				
CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT						1	
VWIAGRYERASQGPASHSRFSRDRG 721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT			1	1		1	
721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYAUIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT		1	J)		
721 2071 A 5632 146 536 MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYAUIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT							VWIAGRYERASQGPASHSRFSRDRG
KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	721	2071	A	5632	146	536	
RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT				3332	1	550	
LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN 722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT]		
722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT				1	1		
722 2072 A 5638 3 3806 CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT		}]	}	J		
TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT				<u> </u>			
LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	722	2072	A	5638	3	3806	
LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT			ł	1			TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES
VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT		1	1	1	i		
GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT			1	}	}		
AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT					ł		
HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT	-				i		
ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT					1		
ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT]	-	
			1	ŀ			ACKHLYPQUACLWACPPGTYQYESWRCVTA
]	}	J	J		
1 1 1 1			1	1			RNSSSIFCHKCEGLCPKECKVGTKTIDSIQAA

NO: of nucleotide sequence NO: of peptide sequence NO: of nucleotide sequence NO: of peptide sequence No: of peptide sequence NEMethionine, N=Asparagine, P=Prol Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptopha y=Tyrosine, X=Unknown, *=Stop cod /=possible nucleotide deletion, \=possible nucleotide insertion QDLVGCTHVEGSLILNLRQGYNLE GLVETITGFLKIKHSFALVSLGFFKI AMVDGNYTLYVLDNQNLQQLGSV PVGKIYFAFNPRLCLEHIYRLEEVT KAEINPRTNGDRAACQTRTLRFVS LLRWERYEPLEARDLLSFIVYYKES HVGPDACGTQSWNLLDVELPLSRT ASLKPWTQYAVFVRAITLTTEEDS	ine, an, on,
sequence uence 09/496 914 orresponding to first amino acid residue of peptide sequence sequence 09/496 914 orresponding to first amino acid residue of peptide sequence sequence sequence M=Methionine, N=Asparagine, P=Prol Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptopha y=Tryrosine, X=Unknown, *=Stop cod /=possible nucleotide deletion, \=possible nucleotide insertion QDLVGCTHVEGSLILNLRQGYNLE GLVETITGFLKIKHSFALVSLGFFKI AMVDGNYTLYVLDNQNLQQLGSV PVGKIYFAFNPRLCLEHIYRLEEVT KAEINPRTNGDRAACQTRTLRFVS LLRWERYEPLEARDLLSFIVYYKES HVGPDACGTQSWNLLDVELPLSR	an, on,
uence 914 ng to first amino acid residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 925 94 127 127 137 137 148 158 158 158 158 158 158 158	an, on,
amino acid residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop cod /=possible nucleotide deletion, \=possible nucleotide insertion QDLVGCTHVEGSLILNLRQGYNLE GLVETITGFLKIKHSFALVSLGFFKI AMVDGNYTLYVLDNQNLQQLGSV PVGKIYFAFNPRLCLEHIYRLEEVT KAEINPRTNGDRAACQTRTLRFVSLLRWERYEPLEARDLLSFIVYYKESLEPLEARDLSFIVYYKESLEPLEARDLSFIVY	on,
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop cod /=possible nucleotide deletion, \=possible nucleotide insertion QDLVGCTHVEGSLILNLRQGYNLE GLVETITGFLKIKHSFALVSLGFFKI AMVDGNYTLYVLDNQNLQQLGSV PVGKIYFAFNPRLCLEHIYRLEEVT KAEINPRTNGDRAACQTRTLRFVSLLRWERYEPLEARDLLSFIVYYKESLURWERYEPLEARDLSFIVYYKESLURWERYEPLEARDLSFIVYYKESL	on,
peptide /=possible nucleotide deletion, \=possible nucleotide insertion QDLVGCTHVEGSLILNLRQGYNLE GLVETITGFLKIKHSFALVSLGFFKI AMVDGNYTLYVLDNQNLQQLGSV PVGKIYFAFNPRLCLEHIYRLEEVT KAEINPRTNGDRAACQTRTLRFVS LLRWERYEPLEARDLLSFIVYYKES HVGPDACGTQSWNLLDVELPLSR3	
sequence nucleotide insertion QDLVGCTHVEGSLILNLRQGYNLE GLVETITGFLKIKHSFALVSLGFFKI AMVDGNYTLYVLDNQNLQQLGSV PVGKIYFAFNPRLCLEHIYRLEEVT KAEINPRTNGDRAACQTRTLRFVS LLRWERYEPLEARDLLSFIVYYKE: HVGPDACGTQSWNLLDVELPLSR3	
QDLVGCTHVEGSLILNLRQGYNLE GLVETITGFLKIKHSFALVSLGFFKI AMVDGNYTLYVLDNQNLQQLGSV PVGKIYFAFNPRLCLEHIYRLEEVT KAEINPRTNGDRAACQTRTLRFVSI LLRWERYEPLEARDLLSFIVYYKESI HVGPDACGTQSWNLLDVELPLSR	
GLVETITGFLKIKHSFALVSLGFFKI AMVDGNYTLYVLDNQNLQQLGSV PVGKIYFAFNPRLCLEHIYRLEEVT KAEINPRTNGDRAACQTRTLRFVS LLRWERYEPLEARDLLSFIVYYKES HVGPDACGTQSWNLLDVELPLSR	POLOHSL
PVGKIYFAFNPRLCLEHIYRLEEVT KAEINPRTNGDRAACQTRTLRFVS LLRWERYEPLEARDLLSFIVYYKE: HVGPDACGTQSWNLLDVELPLSR	
KAEINPRTNGDRAACQTRTLRFVS LLRWERYEPLEARDLLSFIVYYKE: HVGPDACGTQSWNLLDVELPLSR	WVAAGLTI
LLRWERYEPLEARDLLSFIVYYKE: HVGPDACGTQSWNLLDVELPLSR	
HVGPDACGTQSWNLLDVELPLSRT	
PIVYLRTLPAAPTVPQDVISTSNSSS	
KPPTQRNGNLTYYLVLWQRLAED	
YCHRGLRLPTSNNDPRFDGEDGDP	
CCPCQHPPPGQVLPPLEAQEASFQI	
NAITIPISPWKVTSINKSPORDSGRE	
RLGGNSSDFEIQEDKVPRERAVLSO	
RIDIHACNHAAHTVGCSAATFVFA	
ADGIPGKVAWEASSKNSVLLRWLI	
ILKYEIKYRRLGEEATVLCVSRLRY	
HLALLPPGNYSARVRATSLAGNGS YILGPEEEDAGGLHVLLTATPVGL	
LGFFYGKKRNRTLYASVNPEYFSA	
EWEVPREQISITRELGQGSFGMVYE	
AGEESTPVALKTVNELASPRECIEF	
KAFKCHHVVRLLGVVSQGQPTLV	
GDLKSHLRSLRPEAENNPGLPQPA	
AGEIADGMAYLAANKFVHRDLAA	•
DFTVKIGDFGMTRDVYETDYYRKO VRWMAPESLKDGIFTTHSDVWSFO	
TLAEQPYQGLSNEQVLKFVMDGG	
PLQLQELMSRCWQPNPRLRPSFTH	
RPSFRLLSFYYSPECRGARGSLPTT	
TPRDCSPQNGGPGH	'
723 2073 A 5672 1 216 LAWIDNILPEKEKKETDKKRKKK	
EEPQFPPPSVIKIPMESVQSDPQNGI	IHCIARKR
SSSWSYSL	DDDD Ct /
724 2074 A 5704 4235 940 ARGRESPVWAASWGGRGRPAAF	
ATMGFELDRFDGDVDPDLKCALC LTTPCGHVFCAGCVLPWVVQEGS	
LSAKELNHVLPLKRLILKLDIKCAY	
VVKLQQLPEHLERCDFAPARCRHA	
RRDVEAHMRDACDARPVGRCQEC	GCGLPLTH
GEQRAGGHCCARALRAHNGALQA	
ALKKEALRAGKREKSLVAQLAAA	
ALRYQKKFTEYSARLDSLSRCVAA	
ETKSLTLVLHRDSGSLGFNIIGGRP SSSEGIFVSKIVDSGPAAKEGGLOII	
SSSEGIFVSKIVDSGPAAKEGGLQII GRDLSRATHDQAVEAFKTAKEPIV	
PRTKMFTPPSESQLVDTGTQTDITF	
KMSSPSPPVLDPYLLPEEHPSAHEY	YDPNDYI
GDIHQEMDREELELEEVDLYRMN	
VCYRTDDEDDIGIYISEIDPNSIAAK	DGRIREG
DRIIQINGIEVQNREEAVALLTSEE	
ARAELQLDEGWMDDDRNDFLDDI	
EQHHQAMQFTASVLQQKKHDEDO	GEODENIC
ILSNQHEKDSGVGRTDESTRNDES DDATASSNPLAGQRKLTCSQDTLC	
DDATASSNPLAGQREDICSQDTLC NESFISADCTDADYLGIPVDECERF	BELLELK
COVKSATPYGLYYPSGPLDAGKSI	
LELLNEELRSIELECLSIVRAHKMQ	
ESWMLHNSGFRNYNTSIDVRRHEI	LSDITELPE
KSDKDSSSAYNTGESCRSTPLTLEI	SPDNSLRR

	1 200 000	1 > 5	T 0750	10 10 4 1 1 1	D . P. 4. 1 1	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide		in			I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		USSN	location	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	İ	09/496	correspondi	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence	į		914	ng to first amino acid		T=Threonine, V=Valine, W=Tryptophan,
į		ŀ			of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
			l	residue of	sequence	/=possible nucleotide deletion, \=possible
				peptide		nucleotide insertion
	<u> </u>			sequence		
:	1	1				AAEGISCPSSEGAVGTTEAYGPASKNLLSITE
ŀ		ļ	l			DPEVGTPTYSPSLKELDPNQPLESKERRASDG
						SRSPTPSQKLGSAYLPSYHHSPYKHAHIPAHA
1	ľ	1				QHYQSYMQLIQQKSAVEYAQSQMSLVSMCK
]	j	j	ļ			DLSSPTPSEPRMEWKVKIRSDGTRYITKRPVR
[DRLLRERALKIREERSGMTTDDDAVSEMKM
i						GRYWSKEERKQHLVKAKEQRRRREFMMQSR
ļ						LDCLKEQQAADDRKEMNILELSHKKMMKKR
ļ						NKKIFDNWMTIQELLTHGTKSPDGTRVYNSF
						LSVTTV
725	2075	A	5707	3	1770	QISTEVSEAPVANDKPKTLVVKVQKKAADLP
1		1				DRDTWKGRFDFLMSCVGYAIGLGNVWRFPY
1						LCGKNGGGAFLIPYFLTLIFAGVPLFLLECSLG
	1	1				QYTSIGGLGVWKLAPMFKGVGLAAAVLSFW
	1					LNIYYIVIISWAIYYLYNSFTTTLPWKQCDNP
ł						WNTDRCFSNYSMVNTTNMTSAVVEFWERN
1		1			1	MHQMTDGLDKPGQIRWPLAITLAIAWILVYF
1	İ	i				CIWKGVGWTGKVVYFSATYPYIMLIILFFRGV
	1				1	TLPGAKEGILFYITPNFRKLSDSEVWLDAATQ
	1					IFFSYGLGLGSLIALGSYNSFHNNVYRDSIIVC
Ì		ĺ				CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV
ļ.	1					AASGPGLAFLAYPEAVTQLPISPLWAILFFSM
1	1	ĺ		[Ì	LLMLGIDSQFCTVEGFITALVDEYPRLLRNRR
· ·	1	1				ELFIAAVCIISYI IGLSNITQGGIYVFKLFDYYS
		ļ				ASGMSLLFLVFFECVSISWFYGVNRFYDNIQE
1						MVGSRPCIWWKLCWSFFTPIIVAGVFIFSAVQ
1	1		ļ			MTPLTMGNYVFPKWGQGVGWLMALSSMVL
.				1		IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV
1	1)	}		}	RPENGPEQPQAGSSTSKEAYI
726	2076	A	5711	156	423	PRRDPGRTPELRGSAPRKTGANMPVRRGHVA
						PONTFLGTIIRKFEGONKKFIIANARVONCAII
1	ŀ					YCNDGFCEMTGFSRPDVMQKPCTCD
727	2077	Α	5716	3	274	HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP
''	2077	, · ·	3,10	١		LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL
	1	l			ļ	AWFEKMTCYLQLLFNICLPDVSEE
728	2078	A	5737	1899	649	IQASRASPYPRVKVDFALSCHEDLLAPISEPIE
120	20/8	1 ^	3/3/	1000	ا ا	WKYHSPEEEISLGPACWLWDFLRRSQQAGFL
			1			LPLSGGVDSAATACLIYSMCCQVCEAVRSGN
1		1	1			EEVLADVRTIVNQISYTPQDPRDLCGRILTTC
1		ļ.	1			YMASKNSSQETCTRARELAQQIGSHHISLNID
	1	1			1	PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL
1	1	1]	I	ALQNVQARIRMVLAYLFAQLSLWSRGVHGG
	1					LLVLGSANVDESLLGYLTKYDCSSADINPIGG
			1	i		1
	-	ļ	1	1		ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE
		1	1	1	1	LEPLADGQVSQTDEEDMGMTYAELSVYGKL
1	1	1	1		1	RKVAKMGPYSMFCKLLGMWRHICTPRQVAD
		1				KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE
1			[1		DNRFDLRPFLYNTSWPWQFRCIENQVLQLER
		<u> </u>	L	<u></u>		AEPQSLDGVD
729	2079	A	5741	1	5976	PGCAARLSRARAPGPGAAGAGRKRLADPGPP
		1	1	Ι .		PASRRLRAPGSRPRLAPCTRRAAQPAHARMA
1	1	1	1	I	Į.	PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL
	1	1	J	1	j	LLLLLGAARAGALEIQRRFPSPTPTNNFALDG
1		1	1			AAGTVYLAAVNRLYQLSGANLSLEAEAAVG
c	1	1	ŀ		*	PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL
		1	1	1		QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV
ł		1		1	1	AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS
1		1	1	I	1	TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF
		j		1		PRNRSLEDHRFENTPEIAIRSLDTRGDLAKLFT
					·	

KAILSIREDKPPLAVKYFFDFLEEQAEKRGI PDTLHIWKTNSLPLRFWVNILKNPQFVFDII TDHIDACLSVIAQAFIDACSISDLQLGKDSP KLLYAKEIPEYRKIVQRYYKQIQDMTPLSE MNAHLAEESRKYQNEFNTNVAMAEIYKYA	SEQ ID NO: of nucl-cotide seq-uence	NO: of peptide seq-uence	hod	ID NO: in USSN 09/496 914	beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide dinsertion FDLNPSDDNILKIKQGAKEQHKLGFVSAFLHI SDPPPGAQSYAYLALNSEARAGDKESQARSI LARICLPHGAGGDAKKLTESYIQLGLQCAGG AGRGDLYSRLVSVFPARERLFAVFERPQGSP, ARAAPAALCAFRFADVRAAIRAARTACFVEF APDVVAVLDSVVQFTGPACERKLNIQLQPEG LDCGAAHLQHPLSILQPLKATPVFRAPGLTS AVASVNNYTAVFLGTVNGRLLKINLINESMQ VVSRRVVTVAYGEPVHHVMQFDPADSGYLY LMTSHQMARVKVAACNVHSTCGDCVGAAL AYCGWCALETRCTLQQDCTNSSQQHFWTSASEGPSRCPAMTVLPSEIDVRQEYPGMILQISG LPSLSGMEMACDYGNNIRTVARVPGPAFGHG LAYCNLLPRDQFPPPPPNQDHVTVEMSVRVN GRNIVKANFTTYDCSRTAQVYPHTACTSCLS, QWPCFWCSQQHSCVSNQSRCEASPNPTSPQI CPRTLLSPLAPVPTGGSQNILVPLANTAFFQG AALECSFGLEEIFEAVWVNESVVRCDQVVLFTTRKSQVFPLSQLKGRPAFFLDSPEPMTWWYNCAMGSPDCSQCLGREDLGHLCMWSDG RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSSYVLPLVHSLEPTMGPKAGGTRITHGNDLHVGSELQVLVNDTDPCTELMRTDTS ACTMPEGALPAPVPVCVRFERRGCVHGNLTI WYMQNPVITAISPRSPYSGGRTITVAGERFI MVQNVSMAVHHIGREPTLCKVLNSTLITCPS GALSNASAPVDFFINGRAYADEVAVAEELLL PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGEPLTLVHIVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVG LPITIQVGNFNQTIATLQLGGSETAHVSIVICS' LLLLSVVALFVFCTKSRRAERYWQKTLLQM EMESQLREEIRKGFAELQIDMTDLTKELNRSGIPFLEYKHFVTRTFFPKCSSLYEERGVHDSLT WYMGNSCCAARMPKLMLRRTESVVEKMLTNWMSICN SCLASKLITIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICN SCLASKLITIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICN SCUPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKNTLGVKDLDTEKYPREGVCHPERCYVLPSQUPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKNTHLHWKTNSLPLRFWVNILKNPQFVFDIDE KAILSIREDKPPLAVKYFFPPLEEQAEKRGISI PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDL KAILSIREDKPPLAVKYFFPPLEEQAEKRGISI PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDL KAILSIREDKPPLAVKYFFPPLEEQAEKRGISI PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDL KAILSIREDKPLAVKYPFPTRU
730 2080 A 5744 3 292 QPSPLFHSHLETLQLLRTAQLPEQVSWPWC VANGKGNQRNMGSPQPSLLAFERNLELQII	730	2080	Ā	5744	3	292	MEDNIYECYSEA QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMO LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT
	731	2081	A	5747	1	382	

NO: of nucl- peptide eotide seq- uence uence uence uence uence location loc	quence (A=Alanine C=Cysteine, sid, E=Glutamic Acid, ne, G=Glycine, H=Histidine,
eotide seq- seq- uence USSN location corresponding to last amino memory length location corresponding to last amino memory last amino acid residue quence amino acid of peptide length location corresponding to last amino memory last amino acid residue quence, K memory location corresponding last amino memory last amino acid residue quence, K memory last amino acid residue quence, K memory last amino acid residue quence, K memory last amino acid residue quence, K memory last amino acid residue quence last amino acid residue quence quence last amino acid residue quence la constant amino acid residue quence last amino acid residue acid residu	
seq- uence 09/496 correspondi to last amino M=Methionine uence 914 ng to first acid residue Q=Glutamine, amino acid of peptide T=Threonine,	(=Lysine, L=Leucine,
uence 914 ng to first acid residue Q=Glutamine, amino acid of peptide T=Threonine,	, N=Asparagine, P=Proline,
amino acid of peptide T=Threonine,	R=Arginine, S=Serine,
residue of sequence V=Tyrosine X	V=Valine, W=Tryptophan,
	=Unknown, *=Stop codon,
l	leotide deletion, \=possible
sequence nucleotide inse	
	SGRHWKNFALVPLLREASARD
GEK_	YLRQFSGSLKPEDAEVFKSPAAS
PGKLQAQLP	SPEATAGPLCTRIPNVPPPTPIRP CPSPVRFTSARIPPASRPQTKS
	EGRAPESAGPGPGGDAAETPGL
PPAHSGTLM	MAFRDVTVQLANQNISVSSSTAL
	OTVQAPAEPAAGKAEQGETSGR
	REDASAEDSCAEAGASGAADG
	EEEEETAEVGRGAEAEAGDLEQ
	SAKSGSEASASASKDALQAMILS ASCKSPTLSTDTLRKRLYRIGLN
	IOFLISRGFIPDTPIGVAHFLLQRK
	FLGNSKKOFNRDVLDCVVDEM
	ALRKFOAHIRVOGEAOKVERLIE
	CNPEVVQQFHNPDTIFILAFAIILL
	KPDRKMMLEDFIRNLRGVDDG
	GIYERIQQKELKSNEDHVTYVTK
1 1 1 1 1	TVLSVPHRRLVCCSRLFEVTDV
NKLQKQAAI	HQREVFLFNDLLVILKLCPKKKS
	VGLLGMQFQLFENEYYSHGITLV
	QVLHFCALGSDEMQKFVEDLKE
SIAEVTELEC	QIRIEWELEKQQGTKTLSFKPCGA
	SPTAKR FAALRERPAESTVEVSI
	INSGLGAERGAPVPPPDLQPSPPR
	PPTPPGTLVQCQQIVKVIVLDKPC
	QALSCYTSSSSDSCGSTPLGGPG
SPVKVTHQP	PLPPPPPPYNHPHQFCPPGSLLH
	GOGLEEQIVARDENSWLIDGGTP
	DIDEFPQSGNYETIGGFMMFMLR
	KFAGYKFEVVDIDNYRIDQLLVT
	SPKLPDAKDKEESVA
	AFHTGTSNTTFVVYENTYMNITL
	SPLLRYSFETMAPTGLSSLTVNST
AVPTTPAAF	KSLNLPLQITLSAIMIFILFVSFLG
	YQKAAMRSAINILLASLAFADM
LLAVLNMPF	FALVTILTTRWIFGKFFCRVSAMF
	AILLIISIDRFLIIVQRQDKLNPYR
	/ATSFCVAFPLAVGNPDLQIPSRA
	NPGYQAYVILISLISFFIPFLVILY
SFMGILNTLI	RHNALRIHSYPEGICLSQASKLGL
	MSIDMGFKTRAFTTILILFAVFIVC
	VATFSKHFYYQHNFFEISTWLL
1 1 1 1	LNPLIYYWRIKKFHDACLDMMP
	PGHTKRRIRPSAVYVCGEHRTVV HYRTHTGEKRFSCPLCPKQFSRS
	RHPTYHPDMIEYRGRRRTPRIDPP
	SGSGPGPAPSFTTCL
1 1 . 1 1	TLPELLHMSRPAEDGPSPGALVR
	AEEYFLLKSRSDLMFEKQSERH
	ARPPASSEQAQQELFNELKPAV
	IMRDQNNYNEEKDSWNRVART
	TPVMVVGTAWIFLQGVYNQPPP
QPFPGDPYS'	
	AFPNEYTRMSTSELISELFNDCG
738 2088 A 5881 1 1160 LVVTAITAIL	YENRFNTSKGGELPDRPAGVGV
	I DINIG IN I DIZOODDI DIGITO VO
LLDSSKLCD YSAMWQLA	LTLILKIVITIFTFGMKIPSGLFIPS RLLGVGMEQLAYYHQEWTVFNS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion WCSQGADCITPGLYAMVGAAACLGGVTRMT VSLVVIMFELTGGLEYIVPLMAAAMTSKWVA DALGREGIYDAHIRLNGYPFLEAKEEFAHKTL AMDVMKPRRNDPLLTVLTQDSMTVEDVETII SETTYSGFPVVVSRESQRLVGFVLRRDLIISIE NARKKQDGVVSTSIIYFTEHSPPLPPYTPPTLK LRNILDLSPFTVTDLTPMEIVVDIFRKLGLRQC LVTHNGRLLGIITKKDVLKHIAQMANQDPDSI LFN
739	2089	Α	5892	2	916	TLQLAASVPFFAISLISWWLPESARWLIINGKP DQALQELRKVARINGHKEAKNLTIEVLMSSV KEEVASAKEPRSVLDLFCVPVLRWRSCAMLV VNFSLLISYYGLVFDLQSLGRDIFLLQALFGA VDFLGRATTALLLSFLGRRTIQAGSQAMAGL AILANMLVPQDLQTLRVVFAVLGKGCFGISL TCLTIYKAELFPTPVRMTADGILHTVGRLGA MMGPLILMSRQALPLLPPLLYGVISIASSLVVL FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE AFTVESTSLLEIVALHGAL
740	2090	A	5900	2	426	RPIKTLGIGFHFSVDGVHFLTQREVQNLWKE NLIILDTAKKHGYEVVDTFTITMGRYKEFLQG KCGCHFHEVVKSKLSKEYNFIKMKRSRNHIM GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV CSEILLSRMCANKRTM
741	2091	A	5910	3	412	RMPESTLLIICENGYILEAPLPTIKQEEDDHDV VSYEIKDMCIKCFHFSSVKSKILRLIEIEKRER QRELKEKIREERRNKLAAEMGEDGEKEFQEE EEEKEEEEEEEEPLPEIFIPSTPSPILCGFYSEPG KFWV
742	2092	Α	5936	1	482	MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV TQMGNDKSIKCEQNLGHDTMYWYKQDSKK FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL NLHINSLELGDSAVYFCASSQDTALQSHCIPV HKPPGSARKLQGSVCTCTQGSSLHSLMASDG VPVC
743	2093	A	5938	1	1566	MNSFFGTPAASWCLLESDVSSAPDKEAGRER RALSVQQRGGPAWSGSLEWSRQSAGDRRLL GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA ADRARRERFIMNEKWDTNSSENWHPIWNVN DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM CKISGLVQGISVAASVFTLVAIAVDRFQCVVY PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI VALLFILSWLPLWTLMMLSDYADLSPNELQII NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN TSNQLVQESTFQNPHGETLLYRKSAEKPQQE LVMEELKETTNSSEI
744	2094	A	5966	149	327	SHVCVSHYAGSSGCPAGAGAGAVALGISAVA LYDYQGGRLGVARGAWYMEAPDIRQGDM
745	2095	A	5970	413	856	GAPHTDWAWAPTPMSGLGSGRGRQGTLASS PLSLPLLLAGVTGILATELFDQMARPAACMV CGALMWIMLILVGLGFPFIMEALSHFLYVPFL GVCVCGAIYTGLFLPETKGKTFQEISKELHRL NFPRRAQGPTWRSLEVIQSTEL

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl- cotide	peptide seq-		in USSN	location	location corresponding	F=Phenylalanine, G=Glycine, H=Ristidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
		ł		sequence		nucleotide insertion
746	2096	Α	5971	3	1343	AQTARRIIGLELDTEGHRLFVAFSGCIVYLPLS
		ŀ				RCARHGACQRSCLASQDPYCGWHSSRGCVDI RGSGGTDVDQAGNQESMEHGDCQDGATGSQ
i		l				SGPGDSAYGVRRDLPPASASRSVPIPLLLASV
						AAAFALGASVSGLLVSCACRRAHRRRGKDIE
	•]	•	j	TPGLPRPLSLRSLARLHGGGPEPPPPSKDGDA
		i		}		VQTPQLYTTFLPPPEGVPPPELACLPTPESTPE LPVKHLRAAGDPWEWNQNRNNAKEGPGRSR
i						GGHAAGGPAPRVLVRPPPPGCPGQAVEVTTL
					1	EELLRYLHGPQPPRKGAEPPAPLTSRALPPEP
ļ	ļ	l		}		APALLGGPSPRPHECASPLRLDVPPEGRCASA
	ļ		1			PARPALSAPAPRLGVGGGRRLPFSGHRAPPAL
× .		1				LTRVPSGGPSRYSGGPGKHLLYLGRPEGYRG RALKRVDVEKPQLSLKPPLVGPSSRQAVPNG
ł				}	}	GRFNF
747	2097	Α	5998	2	754	DHASLPCSWNHRFDVETRHVFIGDHSGQVTI
1		ļ				LKLEQENCTLVTTFRGHTGGVTALCWDPVQ
						RVLFSGSSDHSVIMWDIGGRKGTAIELQGHN DRVQALSYAQHTRQLISCGGDGGIVVWNMD
	1		l	1		VERQETPEWLDSDSCQKCDQPFFWNFKQMW
						DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL
}						MGFEFEVRVCDSCHEAITDEERAPTATFHDSK
	1	1		[· ·	HNIVHVHFDATRGWLLTSGTDKVIKLWDMT PVVS
748	2098	A	6001	2	747	AMVFGGVVPYVPQYRDIRRTQNADGFSTYV
						CLVLLVANILRILFWFGRRFESPLLWQSAIMIL
]			ļ	}		TMLLMLKLCTEVRVANELNARRRSFTAADS KDEEVKVAPRRSFLDFDPHHFWQWSSFSDYV
		·			j	QCVLAFTGVAGYITYLSIDSALFVETLGFLAV
1						LTEAMLGVPQLYRNHRHQSTEGMSIKMVLM
1	l	l				WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV
749	2099	A	6002	2	447	DLAILGQAYAFARHPQKPAPHAVHPTGTKAL GRPDRSELVRMHILEETFAEPSLQATQMKLK
149	2099	A	0002	-	44 /	RARLADDLNEKIAQRPGPMELVEKNILPVDSS
1						VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP
		1		ŀ		DQPASQESQGSAASPSEPKVSESPSPVTTNTP
750	2100	ļ	6004	 	427	AQFASVSPTVPEFLKTPPTAD LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG
750	2100	Α	6004	2	427	WRWELRLRNYVPEDEDLNKRRVPQAKPDAV
	1	1		1	ĺ	QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD
						WDLKRDVAKKLEKLLKRTQRAIAELIRERLK
761	2101		6007	1 22	1200	GQEDSLDSAVDAATEHKTC
751	2101	Α	6007	33	1280	TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF SHPDKLKRMSKSVPAFLQDESDDRETDTASE
						SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS
				1		VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL
1	1			1		HVFVAQCKDLAAADVKKQRSDPYVKAYLLP DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK
	1			1		QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE
						TWDWDNKQNKQLRWYPLKRKTAPVALEAE
1	ì	1		,		NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV
						KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ
						KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT
	1				٠	EVDWMDSTSEEVALWEKMVNSPNTWIEATL
						PLRMLLIAKISK
752	2102	A	6028	108	1283	KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF
	j]	AAAIPGHRCWVHMLDNNTGSGNETGILSEDA
L	<u> </u>			l	L	JIII I I I I I I I I I I I I I I I

VCVLRFLAGFSSMIIISNNSLPITEWIRPNSKT		·		655	D	D 1	1 4 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
in uncleotide sequence where the corresponding to corresponding to corresponding to corresponding and the sequence where the corresponding to the correspond						_	
uence unce where were pound of the proposed of the proposed of the proposed of the proposed of the proposed of the proposed of the proposed of the proposed of the proposed of the proposed of proposed of the			hod				
equence 99496 corresponding on the first of the second people of peptide sequence continue to the people of peptide sequence continue to the people of peopl							
uence 914 ag to first amino soid of peptide sequence February Februar	eotide	seq-]				
mino soid residue of sequence peptide residue of sequence peptide sequence	seq-	uence	1	09/496	correspondi	to last amino	
residue of peptide pep	uence		f	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
residue of peptide pep			ĺ	ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Popsible nucleotide deletion, "possible nucleotide insertion Incleotide insertion Incleotide insertion Incleotide insertion Incleotide insertion Incleotide insertion Incleotide insertion Incleotide insertion Incleotide insertion Incleotide insertion Incleotide insertion Incleotide Incleotid			1	ł	residue of		
				ĺ			
LLRISPLDSNLRPEKCREFVHPOWQLILLIUG THISTSRADTDEPCUGWVYDQSYPSTIVTKY DLVCDVQSLKSVVQFLLLTOML VGGIIGGIPG SDRFGREIL RWGLIQLAITDTCAAFAPTFPV YCVLRFLAGFSSMIBSNNSLPTEWRPNSKA VVLSSGALNIGQILGGIA VYREDWOYTLHVV ASVPFFVFELLSRWLVESARWLITNKLDEG. KALRKVARTNGKNAEETLNIEVVRSTMQEE LDAAQTKTTVWDLFRNPSMRKRICLIVETRK KALRKVARTNGKNAEETLNIEVVRSTMQEE LDAAQTKTTVWDLFRNPSMRKRICLIVETRK KNLKEKA TOTT VQGSNLNPDLGELVVPLYFKEEKCSI SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGVERRKVYNFRRANSLVMQPISS SVAFTUPGGARATSTANTLAGT GMHTLISTYKAYQQTTLLASSSVEYKKWREPWUF HKFSISENNYIRITSEISPELFMQLVPQYYFSSN PSESKOLLQVVVDLLSPYSTMNKREPWUF HKFSISENNYIRITSEISPELFMQLVPQYYFSSN PSESKOLLQVVDLLSPYSTMNKREPWUF HKFSISENNYIRITSEISPELFMQLVPQYYFSSN PSESKOLLQVVDLLSPYSTMNKREPWUF HKFSISENNYIRITSEISPELFMQLVPQYYFSSN PSESKOLLQVVDLLSPYSTMNKREPWUF HKFSISENNYIRITSEISPELFMQLVPQYYFSSN PSESKOLLQVVDLLSPYSTMNKREPWUF HKFSISENNYIRITSEISPELFMQLVPQYYFSSN PSESKOLLQVVDLLSPYSTMNKREPWUF HKFSISENNYIRITSEISPELFMQLVPQYYFSSN PSESKOLLQVVDLLSPYSTMNKREPWUF HKFSISENNYIRITSEISPELFMQLVPQYYFSSN PSESKOLLQVVDLLSPYSTMNKREPWUF HKFSISENNYIRITSEISPELFMQLVPQYYFSSN PSESKOLLQVVDLLAFTARTATSTANT PSESKOLLGVVDLAARAVFRPSTRCTRANDLUSPYSTMREPWAELLC LGVWALHUFAVPAALLC LGFSSDELWKGRLLPLVLLAGALGFLLPLA VVYSS 755 2105 A 6059 3 1795 LIGGSGTLASKALLYDLAGAKLYCQGV GARARSDTHTFNLFRADEGGRPLLVLTT RAAPGROGRICGGGAAVFVIRESKKTSLVUPAACKKV UFFRASAGRAVLGGLISGGAAVGGRICGGAAVFVIRESKKTSLTV LLTTSVLSSAPADGPRLQGDLRNLCRLARS KYLDGGGTMAAVFRANCLISTSLTLAGATAGAAVGRPLUPALLULL LGGGVTAAVFRANCHTLLTVCGSSIMM ARDLIMNNLTELQFGGFAATAGAAVGRPLUPATALLVLT LGGGGTLASKKKKSLIGHGGAAVGAGAL			ŀ				1 *
THISTSEADTEPCVDGWVPDQSYFFSTINTKS DLVCDVQSLKSVVQFLLTGML VQGIGGIP SDRFGRRFILRWGLLQLAITDTCAARAPIFFEV YCVLRFLAGFSSMISNNSIPTEWRPNSKAI VVLSSGALNIGQIILGGLAVVRDWQTLHVVL ASVPEFVFLLSRWLVSSAWLINTKLDEGL KALRKVARTNGIKNAEFILNIEVVRSTMQEE LDAAQTKITVWDLFRIPSMRRRICILVFLRK KNLKER 753 2103 A 6043 1 1470 DSFESIRLIFHHISGEKGDIVVFLACEQDIER KNLKER TO DSFESIRLIFHHISGEKGDIVVFLACEQDIER VCSTVYQGSNLNPDLGELVVVPLYFKEKCSL FKPLDETERRCQVYQRRVVLTTSSGEFLIWS SVAFVDVGVERRKVYNFRRANSLVMOPISS SQAEIRKQILGSSSSGKFFCLYTEEFASKDMT LRPAEMQEANLTSMYLFMKRDIDAIG,GHCDI MNRPAFESLMQALEDLDVLAALDNDGMLSE FGIMSFPLDPQLSSKLJASCEPTCVEVVLTL AMVTAPNCFSHVPHGAERAALTCWKTFLIFF GGHFTLISIYKAYQDTILNSSEVCVEKWCR YFLNCSALRMADVIRAELLBILKRELEYYAEF FGSKENTINKKALISGYMQIARDVDOSON YLMLTHKQVAQLHPLSCYSTIKKMPEWVLF HKFNISENNYIRTSEISPELFMQLVOYYFSN PSESKDILQQVVDHLSFVSTMNKEQQMCET CPETEGRCTILQ THE STANDAY STANDAY STANDAY STANDAY PSESKDILQQVVDHLSFVSTMNKEQQMCET CPETEGRCTILQ TO THE STANDAY STANDAY STANDAY STANDAY PSESKDILLQVVDHLSRVSTKMPEWVLF LCFESSDELWKGRLLPLVLAEALGFLEPA TO THE STANDAY STANDAY STANDAY STANDAY PSESKDILQVVDHLSRVSTKMPEWVLF LCFESSDELWKGRLLPLVLAEALGFLEPA TO THE STANDAY STANDAY STANDAY STANDAY PSESKDILQVVDHLSRVSTKMPEWVLF LCFESSDELWKGRLLPLVLAEALGFLEPA TO THE STANDAY STANDAY STANDAY STANDAY PSESKDILQVVDHLSRVSTKMPEWVLF LCFESSDELWKGRLLPLVLAEALGFLEPA TO THE STANDAY STANDAY STANDAY STANDAY PSESKDILQVVDHLSRVSTKMPEWVLF LCFESSDELWKGRLLPLVLAEALGFLEPA TO THE STANDAY STANDAY STANDAY STANDAY PSESKDILQVVDHLSRVSTKMPEWVLF LCFESSDELWKGRLLPLVLAEALGFLEPA TO THE STANDAY STANDAY STANDAY STANDAY THE STANDAY STANDAY STANDAY STANDAY THE STANDAY STANDAY STANDAY STANDAY THE STANDAY STANDAY STANDAY STANDAY THE STANDAY STANDAY STANDAY STANDAY STANDAY THE STANDAY		ļ	<u> </u>	 	Sequence		
DLVCDYOSLSSVYOFILLTOML VOGIGGED SORFORREIR WOLLQLAITDICAAPTIFEV YCVLRELAGESSMIIISNNSLPTEWIRPINSKA VVLISSGALNIQOILLGGIAAVTREWOYTLHVV ASVPFFVFELLSRWLVESARWLITNKLDEGL KALRKVARTNGKNAEFILNIEVVRSTMQEE LDAAQTKITVWDLJERNPSMRKRICILLVELKK KALRKVARTNGKNAEFILNIEVVRSTMQEE LDAAQTKITVWDLJERNPSMRKRICILLVELKK KALRKVARTNGKNAEFILNIEVVRSTMQEE LDAAQTKITVWDLJERNIVFJKRKCILLVELKEK VCETVYQGSNLNPDLGELVVVPLYPKEKCSL SVEFVDVGVERRKVYNPRIRANSLVMOPISS SQAEIKRGULGSSSSKOKFFCLYTEEFASKDMT LKPAEMQEANLTSMVLFMKRIDIAGLGHODI MNRPAFESLMQALEDLDYLAALDHOGHSEE FGIIMSEFFLDPQLSKSILASCEPDCVDEVLTI- AMVTAPNCSSVYPTARABLEDLTKLELEYVAEP- FGIMSEFFLDPQLSKSILASCEPDCVDEVLTI- AMVTAPNCSSVYPTARABLEDLTKLELEYVAEP- FGSKENTLINKKALLSGYPMOLARDVDGSSON YLLMITIKQVAQLHPLSGYSTIKKAPEWVLP- HKFSISENNYIRTSEISPELFMQLVPCYYFSNI PPSESKOLQQVVDHLSPVSTIANKEQQNCET- FGSKENTLINKKALLSGYPMOLARDVDGSSON YLLMITIKQVAQLHPLSGYSTIKKAPEWVLP- HKFSISENNYIRTSEISPELFMQLVPCYYFSNI PPSESKOLQQVVDHLSPVSTIANKEQQNCET- FGSKENTLINKAALLSGYPMOLARDVDGSSON YLLMITIKQVAQLHPLSGYSTIANKEQQNCET- FGSKENTLINKAALLSGYPMOLARDVDGSSON YLLMITIKQVAQLHPLSGYSTIANKEQQNCET- GOBSTENSTON YLLTICART STANDARD YLLTICART SINGAPTICART				ļ			
SDRFGRFLRWGLLQLATIDTCAAPAPTIEWRPNSKAI VVILSSGALNIGQIILGGLAYVFRDWOTLH-WASVPFVFFLISKENSKNISNNSLPITEWRPNSKAI VVILSSGALNIGQIILGGLAYVFRDWOTLH-WASVPFVFFLISKENSKVILTSKLDEGL KALKKVARTNGKNAEETLNIEVVRSTMQEE LDAAQTKITTWUDFRNPSMRKRICILVFLRK KNLKEKA Total							
VCVLRFLAGFSSMIIISNNSLPITEWIRPNSKT		ł	ł		i		DLVCDYQSLKSVVQFLLLTGMLVGGIIGGHV
VVILSSGALNIGQIILGGI.AYVFRDWOTH.KLDEGL					ĺ		SDRFGRRFILRWGLLQLAITDTCAAFAPTFPV
ASVPFFVFLLSRWLVESARWLITINKIDES			[1	1		YCVLRFLAGFSSMIIISNNSLPITEWIRPNSKAL
ASVPFFVFLLSRWLVESARWLITINKIDES							VVILSSGALNIGOIILGGLAYVFRDWOTLHVV
RALRKVARTNGIKNAEETLNIEVVRSTMQEE LDAAQTKTITVMUFRIPSMRKRICILVFLRK KNLKEKA Topical London Londo Londo Lon			1	1			
DAAGTKTTVWDLFRNPSMRKRICILVFLRK NNLKEKA 753 2103 A 6043 1 1470 DSFESIRLIFEIHISGEKGEIVVFLACEODIEK VCETVYQGSNLNPDLGELVVVPLYEKKCSL FKPLDETEKRCQVYQRSVLVVITISGGEFLUSS SVAFVIDVGVERRKVYNPRIRANSLVMQPISK SVRFVIDVGVERRKVYNPRIRANSLVMQPISK SVAFVIDVGVERRKVYNPRIRANSLVMQPISK SVAFVIDVGVERRKVYNPRIRANSLVMQPISK SVAFVIDVGVERRKVYNPRIRANSLVMQPISK SVAFVIDVGVERRKVYNPRIRANSLANDGNLSE FKPLDETEKRCQVYQRRVVLTALDNDGNLSE FKPLDETEKRCQVYQRRVVLTALDNDGNLSE FKPLDETEKRCQVYQRRVVLTALDNDGNLSE FKPLDETEKRCQVYQRTANDLADDGNLSE FKPLDETEKRCQVYQRTANDLADDGNLSE FKPLDETEKRCQVTSTRAMPROVEDVLTAL AMMTAPNCFSIVPHGABEAALTCWKTFLIPP GDHFTILISIYKAYQDTTLNSSSEYCVEKWCRI YFLNCSALRMADVIRAELLEIIKRIELPYAEP/FOSKENTLNIKKALLSGYTMQIARDVDGSGN YTMLHKQVAQLHPLSGYSITKKMPEQVID THKPSISENNYINTISSISPELQVPQYYFSNI PHSSISDILQQVVDHLSPVSTIMNKEQQMCET CPETEQRCTLQ 754 2104 A 6055 2 394 YYALHHWPPPDLCQTTGAFQMNMYGSGI LMLNVVDRYAAVHPPLRIKRFRVARILC LGVWALILVFAVPAARVHRPSRCRYRDLEVI LCFES*SDELWKGRLIPLVLLARALGFLIPARVK CHANNAV				}			*
SNLKÉKA SOFESILRIFEIHISGEKGDIVVFLACEODĪTEK VCETVYQGSNLNPDLGELVVVPLYPKEKCSL FKPLDETEKRCQVYQRSVVLTITSGEGERIUWS SVERVIDVOGRKVVTINFRANSLVMQPISG SQAEIRKQILGSSSSGKFCLYTEEFASKOMTI LKPAEMQEANLTISMVLTRINDIAGLGHCDI MINEPAPESLANLTISMVLTRINDIAGLGHCDI MINEPAPESLANLTISMVLTRINDIAGLGHCDI MINEPAPESLANLTISMVLTRINDIAGLGHCDI MINEPAPESLANLTISMVLTRINDIAGLGHCDI MINEPAPESLANLTISMVLTRINDIAGLGHCDI MINEPAPESLANLTISMVLTRINDIAGLGHCDI MINEPAPESLANLAVIRGALEILKRIELEPYAEFF FGIMESPEPLDPQLSVSILASCEPDCVDEVLTI. AMVTAPNCFSHVPHGAEEAALTCWKTFLHP GDHFTILISYKAYQDTTINSSSEYCKEWCRI YFLNCSALTRANDVIRAELEIKRIELPYAEFF FGSKENTLNIKKALLSGYFMQLARVDGSON YLMLTHKQVQQLPHLSGYSTITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYFSNI PPSESKDILQQVVDHLSPVSTIMNKEQQMCET CPETEQRCTLQ 754 2104 A 6055 2 394 YYALHHWPFPDLLCQTTGAIFQMINMYGSCIL LMLINVDRYAAIVHPIRLRHLRRRRVARLLC LGVWALILVFAVPAARVHPSRCRYRDLEVI LCFESFSDELWKGRLLPLVLLAEALGFLLPJA AVVYSS 141 LAVYSS LAVYSS LAVYSS LAVYSS LAVYSS LAVYSS RAVSSYKTKRPTKLLIAPPSAAPEEALGPAE PPEPGARRSDTHTFNRLFRRDEGGRRPLTVV QQPAGIGKTMAAKKLLYDWAAGKLYQOQW FAFFMPCGELLERPGTRISLADLLIDQCPBGG PPYQMLAQPQRALFILDADAELPALGPAE PPEPGARRSDTHTFNRLFRRDEGGRRPLTVV QQPAGIGKTMAAKKLLYDWAAGKLYQOQW FAFFMPCGELLERPGTRISLADLLIDQCPBGG PPYQMLAQPQRLIFILDADAELPALGPEA CTIPFEAASGARVLGGILSKALLFTALLIVT RAAAGGRILQGILSPACAETISTSTTTSV LIFTSVLSSAPVADGPRLQGLRNIKTTTSV LIFTSVLSSAPVADGPRLQGLRNIKTTTSV LIFTSVLSSAPVADGPRLQGLRNIKTTTSV LIFTSVLSSAPVADGPRLQGLRNIKTTTSV LIFTSVLSSAPVADGPRLQGLRNIKTTTSV LIFTSVLSSAPVADGPRLQGLRNIKTTTSV LIFTSVLSSAPVADGPRLQALGCRIFEK LIRUVQQGQQCCPGVAPEVTEGAKGLEDTE PPEEBEGEEPTYP ELLV LYTCPGDAFVRQ LRWVQQGQQCCPGVAPEVTEGAKGLEDTE PPEEBEGEEPTYP ELLV LYTCPGDAFVRQ LCRFPELALQRVRFCMDVAVLSYCVRCCP QQALRISCRLVAAQEKKKSLGRRLQSIM LWUDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEIPS LWLDDNALTEI			l			į	
Total							, -
VCETVYQGSNLNPDLGELVVVPLYVEKECSI FEPI.DETEKRQVVQRRVVJTTSGGEFLIWS SVRFVIDVGVERRKVYNPRIRANSLVMQPISG SQAEIKRQLGSSSSGKFFCLYTEEFASKDMTI LKPPAEMQEANLTSMVLTMKRIDIAGLGHCDI MRPAPESLMQALEDLDYLAALDNDGRLSE FGIIMSEPPLDPQLSKSILASSFVCFWEWGI AMPTAPESLMQALEDLDYLAALDNDGRLSE FGIIMSEPPLDPQLSKSILASSFVCFWEWGI OFFTLISIYRAYQDTTLNSSFVCFWEWGI FOSKENTLNIKKALLSGYFMQLARDVDGSGN YLMLTHKQVQQLHPLSGYSTIKKMEPWVLF HKFSISENNYIRITSEISPELFMQLVPQYFSNI PPSESKDILQQVVDHLSPVSTIMNEQQMYCFS TOSKENTLNIKKALLSGYFMQLARDVDGSGN YLMLTHKQVQQVDHLSPSTMINEQQMYCFS PPSESKDILQQVVDHLSPSTMINEQQMYCFS CPETEQRCTLQ TYALHHWPPPDLLCQTTGAIPQMNMYGSCI LMLINVDRYAAIVHPLRLHLRRPRVARLLC LGVWALILVFAVPAARVHPSRCRYRDLEVI LCFESSPBELWGRULPLVLLAEALGFLLPA AVVYSS TOSKENTLANKKALLSGYFMQLARDVSTMINEQQMYCFS PPQMLAQPUBLAGFEARARVFYSLENTERPSRCRYRLTVY QGPAGIGKTMAAKKLLYDWAARKLYQQVI FAFFMPGGELLERPGTRSILADLLDQCPBRGA PPPQMLAQPGRLIFILDGADELPALGPEA PEPGARRSDTHTFIRLIFREGERRPLTVY QGPAGIGKTMAAKKLLYDWAAGKLYQQVI FAFFMPGGELLERPGTRSILADLLDQCPBRGA PPQMLAQPGLIFLDGADELPALGPEA CTDFFEAASGARVLGGLLSKALLFTALLVT RAAAFGRLQGRLCSPQCAEFGSKDKKK VYFYKFFRDERRABRAYFFVKENETLFALCFV PPVCWIVCTVLRQQLELGRGRSTMITLITVT RAAAFGRLQGRLCSPQCAEFGSKDKKK VYFYKFFRDERRABRAYFFVKENETLFALCFV PPVCWIVCTVLRQQLELGRGRSTMITLITVT RRITGILSABRIMDIERFIEGGMVSERVKQEA LFFTSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFASKELEQLELRGSKVOTIFLSK KELPGVLETEVTYQFIDOSFGVELAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHVLJTT RRITGILSABRIMDIERFIEGGMVSERVKQEA LFFTSVLSSAPVADGRKKKSLGRRLQASLC GG T56 2106 A 6060 12 436 SGRPTRPAKPTQQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEERLSGNN LSHIPGQAPSGLYSLKILMLHNNQLGGIMA ARDLIMNNLTELQPGLFHLIRFLEERLSGNN LSHIPGQAPSGLYSLKILMLHNNQLGGIMA ARDLIMNNLTELQPGLFHLIRFLEERLSGNN LSHIPGQAPSGLYSLKILMLHNNQLGGIMA ARDLIMNNLTELQPGLFHHLRFLEERLSGNN LSHIPGQAPSGLYSLKILMLHNNQLGGIMA ARDLIMNNLTELQPGLFHHLRFLEERLSGNN LSHIPGQAPSGLYSLKILMLHNNQLGGIMA ARDLIMNNLTELQPGLFHHLRFLEERLSGNN LSHIPGQAPSGLYSLKILMLHNNQLGGIMA ARDLIMNSLTELPPEFEVENI THILGGAADMCAFFWILLLLLLLQEGGRRL WWQGSECVAVALQESISLSLIJLEIPPPEFEVENI THILGGAADMCAFFWILLLLLLLQEGGRRL			<u> </u>				<u></u>
FEPLDETEKRCQVYQRRVVLTYSGEPLIWSS SVRFVDVGVERRXVYPRIRANSLVMQPISK SQAEIRKQLIGSSSSGKFFCLYTEEFASKDMTI LKPAEMQEANLTSMVLFMKRIDIAGLGHCD MNRPAPESLMQALEDLDYLAALDNDGNLSE FGIIMSEFPLDPQLSKSILASCEFDCVDEVLTIL AMVTAPNCPSHVPHGAEALTCWKTFLIPP GDIETILISIYKAYQDTTLNSSSEYCVEKWCRI YFLNCSALRMADVIRAELLEIIKRIELPYAEPP FGSKENTLNIKKALLSGYBIQLARDVOGSGN YLMLTIKQVAQLIFILSGYSITKKMEWWLF HKFSISENDILQQVVDHLSPVSTMIKEQQMCET CPETEGORCILQ 754 2104 A 6055 2 394 YYALHHWPFPDLLCQTTGAIFQMNMYGSCII LMLINVDRYAAVYHPLRIRHLRRFWARLC CGYWALLIVFAVPAARVHEPSRCEYRDLEVI LCFESFSDELWKGRILPLVLLAEALGFLLPLA AVVYSS 755 2105 A 6059 3 1795 LGIGSGTLLSVSEYKKKYREHVOLHARVK RNARSVKITKRPTKLLIAPESAAPEEALGPAE PEPGRARSDTHTINRLFRIDEEGGRPLIVVI QGPAGIGKTMAAKKILYDWAAGKLYQGQVI FAFFMPCGELLERPGTSALDILDQCPDRG- PVPQMLAQPQRLIFILDGADELPALGPEAA CTDPFEAASGARVI.GGLISKALLPTALLLUT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YYYKFFRDERRAERAYRFVKENETI-FALCTV PPVCWIVCTVLRQQLEJGRDLSRTSKTITISV LLFITISVLSSAPVADGPRLQGDLARNICRI.ARR GVI.GRRAGGARVI.GGLISKALLPTALLLLUT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YYYKFFRDERRAERAYRFVKENETI-FALCTV PPVCWIVCTVLRQQLEJGRDLSRTSKTITISV LLFITISVLSSAPVADGPRLQGDLRGKKKKSLIPTALLLUT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YYYKFFRDERRAERAYRFVKENETI-FALCTV PPVCWIVCTVLRQQLEJGRDLSRTSKTITISV LLFITISVLSSAPVADGPRLQGDLRGKKVCTIFLSK KELPGVLETEVTYQFDQSSCPET-AALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLIT RFLFGLISAERMRDIERHFGCMVSERVKQEA LRWVQGQQGCGCFGVAPPVTEGAKCLEDTEE PEEEEGGEFENTYPLELLYCLYTETQEDAPFRQ- LCRFPELALQRVRFCRMDVAVLSYCVRCCP- GQALRLISCRLVAAQEKKKKSLGRRLQSLVC LGGG 756 2106 A 6060 12 436 SGRPTPAKPTGQGMGRFMLTLVCQSIMM ARDLIMNNLTELQPGLFHHLRFLEELR.SGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQ- LWELPSLOSRLAIANLISLVPERSFEGLSLRI LWLDDNALTEIPS 117PLGLGAADMCAFFWLLLLLLQEGSQRRL VWCCSSEEVVAVLQESISLPLEIPPDEEVENI	753	2103	Α	6043	1	1470	DSFESILRLIFEIHHSGEKGDIVVFLACEQDIEK
SVRFVIDVGVERRYVPHERARMLVMQPISS SQAEIRKQILGSSSSGKFFCLYTEEFASKDMT LKPAEMQEANLTSMVLFMKRIDIAGLGHCDI MNRRAPESLMQALEDLDYLAALDNDGNLSE FGIIMSEFPLDPQLSKSLAGEFDCVDEVLTIL AMVTAPNCFSHVPHGAEEAALTCWKTFLIPP GBHFTLISIYKAYQDTHSSSEYCVEKWCRI YFLNCSALRMADVIRAELLEIIKRIELPYAEPAFOSKENTLNIKKAALLSGYFMQIARDVDGSGN YLMLTIKQVAQLIPHZSGYSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYYFSNI PFSESKDLQQVVDHLSPVSTMIKEQQMCET CPETEQRCTLQ 754 2104 A 6055 2 394 YYALHTWYPFDLLCQTTGAIFQMNMYGSCII LMLINVDBYAAIVHPLRLRHLRRPRVARLLC LGVWALLJVFAVPAARVHRPSRCRYRDLEVI LCFESFSDELWKGRLLPLVLLAEALGFLLPLAVVYSS 755 2105 A 6059 3 1795 LGLGSGTLLSVSEYKKKYREHVLQHARVKI QAPAGIGKTMAAKKILYDWAAGKLYQGQV QAPAGIGKTMAAKKILYDWAAGKLYQGQV GAPAGIGKTMAAKKILYDWAAGKLYQGVV QAPAGIGKTMAAKKILYDWAAGKLYQGVV QAPAGIGKTMAAKKILYDWAAGKLYQGVV FAFFMPCGELLERPGTRSLADLLDQCPDRGAPPPQMLAQPQRLLFILDGADELPALGGPEAA CTTDFFEAASGARVLIGGLLSKALLFTALLLUT RAAAPGRLQGRLCSPQCAEVRGPSDKDKKK YFYKFFRDERRAERAYRFVKNETLFALCTV PFVCWIVCTVLRQQLELGGDLSRTSKTTTSV LLFITSVLSSAPVADGPRLQGDLRNLCRLARG GVLGRRAQFAEKLEQLELRGGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQETAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQQQGQCPGVAPEVTEGAKGLEDTE PEEEEGGEENYTPLELLYCLYTETQEDAFVRQ. LCRFFELALQRVRFCRMDVAVLSYCVRCCP-GQALRISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRAKPTQGMGFFMLTLVCQGSIMM ARDLIMNNLTELQPGLPHHLRFLEELR.SGNI LSHIPGQAFSGLVSLKILMLHNNJCGGIPAQA LWELPSLQSLRLDANLSLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLQEGSGRRI					1	1	VCETVYQGSNLNPDLGELVVVPLYPKEKCSL
SVRFVIDVGVERRYVPHERARMLVMQPISS SQAEIRKQILGSSSSGKFFCLYTEEFASKDMT LKPAEMQEANLTSMVLFMKRIDIAGLGHCDI MNRRAPESLMQALEDLDYLAALDNDGNLSE FGIIMSEFPLDPQLSKSLAGEFDCVDEVLTIL AMVTAPNCFSHVPHGAEEAALTCWKTFLIPP GBHFTLISIYKAYQDTHSSSEYCVEKWCRI YFLNCSALRMADVIRAELLEIIKRIELPYAEPAFOSKENTLNIKKAALLSGYFMQIARDVDGSGN YLMLTIKQVAQLIPHZSGYSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYYFSNI PFSESKDLQQVVDHLSPVSTMIKEQQMCET CPETEQRCTLQ 754 2104 A 6055 2 394 YYALHTWYPFDLLCQTTGAIFQMNMYGSCII LMLINVDBYAAIVHPLRLRHLRRPRVARLLC LGVWALLJVFAVPAARVHRPSRCRYRDLEVI LCFESFSDELWKGRLLPLVLLAEALGFLLPLAVVYSS 755 2105 A 6059 3 1795 LGLGSGTLLSVSEYKKKYREHVLQHARVKI QAPAGIGKTMAAKKILYDWAAGKLYQGQV QAPAGIGKTMAAKKILYDWAAGKLYQGQV GAPAGIGKTMAAKKILYDWAAGKLYQGVV QAPAGIGKTMAAKKILYDWAAGKLYQGVV QAPAGIGKTMAAKKILYDWAAGKLYQGVV FAFFMPCGELLERPGTRSLADLLDQCPDRGAPPPQMLAQPQRLLFILDGADELPALGGPEAA CTTDFFEAASGARVLIGGLLSKALLFTALLLUT RAAAPGRLQGRLCSPQCAEVRGPSDKDKKK YFYKFFRDERRAERAYRFVKNETLFALCTV PFVCWIVCTVLRQQLELGGDLSRTSKTTTSV LLFITSVLSSAPVADGPRLQGDLRNLCRLARG GVLGRRAQFAEKLEQLELRGGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQETAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQQQGQCPGVAPEVTEGAKGLEDTE PEEEEGGEENYTPLELLYCLYTETQEDAFVRQ. LCRFFELALQRVRFCRMDVAVLSYCVRCCP-GQALRISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRAKPTQGMGFFMLTLVCQGSIMM ARDLIMNNLTELQPGLPHHLRFLEELR.SGNI LSHIPGQAFSGLVSLKILMLHNNJCGGIPAQA LWELPSLQSLRLDANLSLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLQEGSGRRI				1	İ	[FKPLDETEKRCQVYQRRVVLTTSSGEFLIWSN
SQAEIRKQILGSSSSKFFCLYTEEPASKDMT LKPAEMQEANLTSMVLRRIDIAGLGHCD MNRPAPESLMQALEDLDYLAALDNDGNLSE FGIIMSEFFLDPQLSKSILASCEFPCVEVEVLTIL AMYTAPNCSHVPHGAEAALTCWKTFLHP GDHFTLISIYKAYQDTTLNSSSEVCVEWCRG YFLNCSALMADVIRAELLEIIKRIELPYAEPA FGSKENTLNIKKALLSGYFMQIARDVDGSGN YYALTHINVEVAQLHPLSGYSTIKKMPEWVLF HKFSISENNYIRITSEISELFIMQLVPQYYFSN PSESKDILQQVVOHLSPVSTINKKEQQMCET CPETEQRCTLQ TYALHHWPFPDLLCQTTGAIFMMYGSGI LGUWALILVFAVPAARVHPLRLHRLRRPRVARLLC LCVWALILVFAVPAARVHPLRLHRLRRPRVARLLC LCVWALILVFAVPAARVHPRSRCRYRDLEVI LCFSFSDEL WKGRLLPLVLLAEALGFLLPLA A 6059 3 1795 LGLGSGTLLSVSEYKKKYREHVLQLHARVK RNASVSUTIKRPTKLLIAPESAAPEEALGPAE PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVI QGPAGIGKTMAAKKILYDWAAGKLYQGOV FAFFMPCGELLERPGTSILADLLDQCPDRG PVPQMLAQPQRLLFILDGADELPAGPEA CTOPFFAASGARVI.GGLISKALLFILALLLV RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRARRAYRFVKENETLFALCFV PPVCWVCTVLRQQLELGRDLSRTSKTITSV LLFITSVLSSAPVADGPRLQGDLSRTSKTITSV LLFITSVLSSAPVADGPRLQGDLSRTSKTITSV LLFITSVLSSAPVADGPRLQGDLSRTSKTITSV LLFITSVLSSAPVADGPRLQGDAPHSH VLTT RAAAPGRLAGRGAVGGAQPHSH VLTT RFLFGLLSAERMRDIERHFGCMVSERVKQE LRWVQGQGGCFGVAPPVTEGAKGLEDTE PEEEEEGEFERPTYPLELLYCLYETQEDAPVRQ LCRFFELALQRVRFCRMDVAVLSYCVRCCP GQALRLISCRLVAAQEKKKSLGKRLQASLC GG					1		SVRFVIDVGVERRKVYNPRIRANSLVMOPISO
LKPAEMOEANLTSMYLPMKRIDIAGLGHCDI MNRPAPSLMQALEDLDYLAALDNDGNLSE FGIIMSEPFLDPQLSKSILASCEFDCVDEVLTI/ AMVTAPNCFSHVPHGAEEAALTCWKTFLHP GDHFTLISIYKAYQDTTLNSSSEYCVERWCRI YFLNCSALRMADVIRAELLEIKRIELPYAEPP FGSKENTLNIKKALJSGIPBIKRIELPYAEPP FGSKENTLNIKKALJSGIPBIKRIELPYAEPP FGSKENTLNIKKALJSGIPBIKRIELPYAEPP FGSKENTLNIKKALJSGIPBIKRIELPYAEPP FGSKENTLNIKKALJSGIPBIKRIELPYAEPP FGSKENTLNIKKALJSGIPBIKRIELPYAEPP FGSKENTLNIKKALJSGIPBIKRIELPYAEPP FGSKENTLNIKKALJSGIPBIKRIELPYAEPP FGSKENTLNIKKALJSGIPBIKRIELPYAEPP FGSKENTLNIKTRISSISPELFMQLVOYYSSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVOYYSOK PEFEGRCTLQ YYALHHWPFPDLLCQTTGAIPQMMMYGSCII LGWWALILVFAVPAARYHRPSRCRYRDLEVY LCFESSPELWKGRLPLVLLAEALGFILLPA LCFESSPELWKGRLPLVLLAEALGFILLPA LCFESSPELWKGRLPLVLLAEALGFILLPA LCFESSPELWKGRLPLVLLAEALGFILLPA RNARSVKITKRFTKLLIAPSSAPEEALGPAE PEFGRARRSDITHTYNELFREDEGRRPLT-VVY QGPAGIGKTMAAKKLVDWAAGGLLYQGVY FAFFMPGGELLERPGTRSLADLILDQCPDRG4 PVPQMLAQPQRLLFILDGADELPALGGPEAA CTDFFBAASGARVLGGLISKALLTPALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFRDERRAERAFVYKENETI-FALCTV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSVY LLFITSVLSSAPVADGPRLQGDLRNLCRLAR GVLGRRAQFAEKELGQLELRGSKVQTLFLSK KELPGVLETSVTYQFDQSFGCFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDERHFGCMVSERVKQEA LRWVQGQGQCPGVAPVPTUFEGAKGLEDTEF PEEBEEGEEPNYPLELLYCLVETGEDAFVRQ, LCREPFELALQRVBFCRMDVAVLSVCVRCCP, GQALRLISCRLVAAQEKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQ, LWELPSLQSLRLDANLTVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQ, LWELPSLQSLRLDANLTVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQ LWELPSLQSLRLDANLTVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQ LWELPSLQSLRLDANLTVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILLQESGSRRL LWLDDNALTEPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQESGRRL WRWCSSEEVVAVLQESISLLEEPPDEEVENI		1	1	1	1	[
MNRPAPESLMQALEDLDYLAALDNDGNLSS FGIIMSEFPLDPQLSKSILASCEFDCVDEVLTI, AMVTAPNCFSHVPHGAEEAALTCWKTFLHP GDHFTLISIYKAYQDTTLNSSSEYCVEKWCR YFLNCSALRMADVIRALEIHKRELEJYAEPA FGSKENTLNIKKALLSGYPMQIARDVDGSGN YLMLTHKQVAQLHPLSGYSITKKMPEWVLF HKFSISENNYRITSEISPELFMQLVPQYYFSNI PPSESKDILQQVVDHLSPVSTMNKEQQMCET CPETEGRCTLQ 754 2104 A 6055 2 394 YYALHHWPFPDLLCQTTGAIFQMNMYGSCII LMLINNVDRYAANVHPLRLRHLRPRVARLLC LGVWALILVFAVPAARVHRPSRCRYRDLEVI LCFESFSDELWKGRLLPLVLAEALGFLLPLA AVVYSS 755 2105 A 6059 3 1795 LGLGSGTLLSVSEYKKKYREHVLQLHARVKI RNARSVKITKRFTKLLIAPESAAPEEALGFALE PEGRARSDTHTFINFIRLFRDEEGREPLTVV QGPAGIGKTMAAKKILVDWAAGKLYQGQVV FAFFMPCGELLERPGTRSLADLILDQCPDRG/ PVPQMLAQPQRLLFILDGADELPALGGPEAA CTDPFEAASGARVJGGLISKALLPTALLLLVT RAAAPGRLQGRLCSPQCAEVRGFSBOKDKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRISKTITISV LLFITSVLSSAPVADGPRLQGDLRNICRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWQGQGGCGCPAVSTEFSKGLEDTE PEEEEGGEEPNYPLELLVCLYETQEDAFVRQC, LCRFPELALQRVRFCRMDVALYSVCVRCCP GQALRLISCRLVAAQEKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRRMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRSON LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDAAQEKKKSLGKRLQASLC GG 757 2107 A 6063 54 419 ITPLGLGAADMCAFFWLLLLLLLQESGQRRL WRWCGSEEVVAVLQESISLPLEPPDEEVENI	1	1	l	1		Į	
FGIIMSEPPLDPQLSKSILASCEPDCVDEVITLI AMVTAPNCFSHVPHGAEEAALTCWKTFLHP GDHFTLISIYKAYQDTTLNSSSEVCVEKWCRI YFLNCSALRMADVIRAELLEIHKRIELPYAEPA FGSKENTLNIKKALLSGYPMQIARDVDGSON YLMILTHKQVAQLHPLSGYSITKKMPEWLIF, HKFSISENNYIRITSEISPELFMQLVPQYYFSNI PSESKDILQQVDHLSPVSTMNKEQQMCET CPETEGRCTLQ 754 2104 A 6055 2 394 YYALHHWPPPDLCQTTGAIFQMMMYGSGI LMLNNVDRYAAIVHPLRLRIERRPRVARLLC LGVWALILVFAVPAARVHPLRRIERRPRVARLLC LGVWALILVFAVPAARVHPLRCHLRRPRVARLLC LGGVWALILVFAVPAARVHPLSPSRCRYRDLEVI LCFESSSDELWKGRILPLV/LLAEALGFLLPLA AVVYSS 755 2105 A 6059 3 1795 LGIGSGTLLSVSEYKKKYREHVLQLHARVK RNARSVKITKRPTKLLIAPESAAPEEALGPAE PEPGRARRSDTHTFINELFREDEGRRPLT-VVI QGPAGIGKTMAAKKLYDWAAGKLYQQQVI FAFFMPGGELLERPGTRSLADILDQCPDRG PVPQMLAQPRALIPLDGADELPALGGPEAA CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAPGRLQGRICSPQCAEVRGFSDKDKKK YTYKFFRDERRAERAFNYEKNENT-FALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV LLFITSVLSSAPVAGREQGDLCRNLCRLAR GVLGRRAQFAEKELEQLERGSKVQTLFLSK KELPGVLETEVTYQFDQSFGETAALSYLLE DGGVPRTAAGGVGTLLGDAQPFBHLVLTT RFLFGLLSAERMDIERHFGCMVSERVKQEA LRWVQGQQGQCDVAPEVTGEAK GLEDTER PEEEEGEEPNYPLELLYCLYETQEDAFVRQ, LCKPPELALQRVBFCRMDVAVLSVCVRCCP GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMRGFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDAILSVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGGAAADMCAFPWLLLLLLLQESGRRL			1				
AMVTAPNCPSHVPHGAEEAALTCWKTFI.HP GDHFTLISIYKAYQDTTLNSSSEYCVEK WCRI YFLNCSALRMADVIRAELLEIIKRIELPYAEP FGSKENTLNIKKALLSGYPMQIARDVDGSON YLMLHKQVAQHPLSGYSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYYFSNI PPSESKDILQVVDHLSPVSTMNKEQQMCET CPETEQRCTLQ CPETEQRCTLQ T54 2104 A 6055 2 394 YYALHHWPPPDLLCQTTGAIFQMMMYGSCII LMLINVDRYAAIVHPLRLHRLRPRVARLLC LGVWALLLVFAVPAARVHRPSRCRYRDLEVI LCFESFSDELWKGRI.LPLVLLAEALGFLLPLA AVVYSS T55 2105 A 6059 3 1795 LGLGSGTLLSVSEYKKKYREHVLQLHARVKI RNARSVKITKPTKLLAIPESAAPEEALGFAE PEPGRARRSDHTHRIFRRDEEGRPLTVVI QGPAGIGKTMAAKKILYDWAAGKLYQGQVI FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLIFILDGADELPALGGFEAA CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAFGRLQGRLSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCEV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYPQDSFOEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERRIFGCMYSERVKQEA LRWVQGQGQGCTPDOSFOEFLAALSYLLE LGVURETEVTYPLOSFOEFLAALSYLLE LGVURETEVTYPLOSFOEFLAALSYLLE GGG T56 2106 A 6060 12 436 GGPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHLPQQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSILADANISLVPERSFEGLSSLRI LWLDDNALTEIPS T757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLLQEGSQRRL		1		1			
GDHFTLISIYKAYQDTTLNSSSEVCVEKWCRI YFLNCSALRMADVITAAGLLEIIKRIELPYAEPA FGSKENTLNIKKALLSGYFMQLARDVDGSGN YLMLTHKQVAQLHPLSGYSTIKKMPEWVLF HKFSISENNYIRITSEISPELFMQL VPQYYFSNI PPSESKDILQQVVDHLSPVSTMNKEQQMCET CPETEQRCTILQ 754 2104 A 6055 2 394 YYALHHWPPPDLLCQTTGAIFQMNMYGSGI LMLINVDRYAAIVHPLRLHRLRRPRVARLLC LGVWALILVFAVPAARVHRPSRCRYRDLEVI LCFESFSDEL WKGRILPLVLLAEALGFLLPLA AVVYSS 755 2105 A 6059 3 1795 LGLGSGTLLSVSEYKKKYREHVLQLHARVKI RNARSVKITKRPTKLLIAPESAAPEEALGFALE PEPGRARSDTHTFNRLFRDEGGRPLTVVI QGPAGIGKTMAAKKLYDWAAGKLYQGQVI FAFFMPGGELLERPGTBSLADLLDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPBAA CTDPFEAASGAVLGGLLSKALLYTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERAERAYRFYKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTITTSVV LLFITSVLSSAPVADGPRLQGDLRNLCRLARS GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAAGGVGTLIRGAQPHSHLVJT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQQGCCPGVAPEVTGGAKGLEDTEE PEEEEGEGEPNYPLELLIYCTYGTDAFVRQU LCRPFELALQRWRFCRMDVAVLSYCVRCCP- GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPQQAFSGLYSKLMINNNQLGGPAQA LWELPSLQSIRLDANNLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQCGSQRRL		{	ĺ	l	1	ĺ	,
FINCSALRMADVIRAELLEIIKRIELPYAEPA FGSKENTLNIKKALLSGYFMQIARDVDGSGN YLMLTHKQVAQLIPLSGYSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYYPSNI PPSESKDILQQVVDHLSPVSTMNKEQQMCET CPETEQRCTLQ T54 2104 A 6055 2 394 YYALHHWPPDLLCQTTGAIFQMNMYGSGI LMLINVDRYAAIVHPRLRHLRPRPVARLLC LGVWALLLVFAVPAARVHRPSRCRYRDLEVI LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS T55 2105 A 6059 3 1795 LGLGSGTLLSVSEYKKKYREHVLQLHARVKI RNARSVKITKRFTKLLIAPESAAPEEALGFAE PEPGRARSDSTHTRIFRIFREGERRPLTVVV QGPAGIGKTMAAKKILYDWAAGKLYQGQVI FAFFMPCGELLERFGTBSLADLILDQCPDRGA PVPQMLAQPRGLIFILDGADELPALGGPEAA CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAFGRLQGRLCSPQCAEVRGFSDKOKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV: LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELQQLELRGSKYQTLFLSK KELPGVLETEVTYQFDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRDDAQPHSHLVLTT RFLFGLLSAERMDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCP- GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGJFHHLRFLEELRISGNI LSHIPGQAFSGLYSLKILMILHNNQLGGIPAQA LWELPSLQSIRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQGGSQRRL WRWCGSEEVVALQESISLPLEIPPDEEVERL			ł		1		1
FGSKENTLNIKKALL.SGYFMQLARDVDGSGN YLMLTHKQVAQLHPLSGYSITKKMPEWVLF HKFSISENVIJRITSEISPELFMQLVPQYYFSNI PPSESKDILQQVVDHLSPVSTMNKEQQMCET CPETEQRCTLQ 754 2104 A 6055 2 394 YYALHHWPFPDLLCQTTGAIFQMNMYGSCII LMLINVDRYAAIVHPILRIHLRRPRVARLLC LGVWALL VFAVPAARVIRPSRCRYRDLEVI LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS 755 2105 A 6059 3 1795 LGLGSGTILSVSEYKKKYREHVLQLHARVKI RNARSVKITKRFTKLLIAPESAAPEEALGPAE PEPGRARRSDTHITFINEIRRDEEGRPLT-VVI QGPAGIGKTMAAKKILYPWAAGKLYQGQVI FAFFMPGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRILFILDGADELPALGGPEAA CTDPFEAASGARVLGGLISKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLEIGRDLSRTSKTITSV LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKLEGOLELREGSKVOTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMDDIERHGGCMVSERVKQEA LRWVQGQGQCPGVAPEVTEGAKGLEDTEE PEEEEGEEPNYPLELLYCLYETQEDAFVRQL LCRFPELALQRVRFCRMDVAVLSYCVRCQ- GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNITTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSIKILMHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLILLLLLQEGSQRRL	1			1	1	}	GDHFTLISIYKAYQDTTLNSSSEYCVEKWCRD
YIMLTHKQVAQLHPLSGYSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYYFSN PPSESKDILQQVVDHLSPVSTMNKEQQMCET CPETEQRCTLQ 754 2104 A 6055 2 394 YYALHHWPFPDLLCQTTGAIFQMNMYGSCII LMLINVDRYAAVIHPLRLRHLRRPRVARLLC LGVWALILVFAVPAARVHRPSRCRYRDLEVI LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS 755 2105 A 6059 3 1795 LGLGSGTLLSVSEYKKKYREHVLQLHARVKI RNARSVKITKRPTKLLIAPESAAPEEALGPAE PEPGRARRSDIHTHTNRLFRADEEGRRPLTVVI QQPAGIGKTMAAKKILYDWAAGKLYQGQVI FAFFMPCGELLERPGTRSLADLILDQCPDRGA CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKITTSV LLFITSVLSSAPVADGPLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RRLFGLLSAERMDDERHFGCMVSERVQGA LRWVQGQGQCGCGVAPEVTEGAKGLEDTEE PEEEEGEEPNYPLELLYCLYETQEDAFVRQ LCRFPELALQRVRFCRMDVAVLSYCVRCCP GQALRLISCRLVAAQEKKKSLGKRLQASLC GG GG SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELOPGLFHHLRRLEELRLSGN LSHIPQQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLR LWLDDNALTEIPS LWLDDNALTEIPS SSRIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVEN	1					j	YFLNCSALRMADVIRAELLEIIKRIELPYAEPA
HKFSISENNYIRITSEISPLFMQLVPQYYFSNI PPSESKDILQQVVDHLSPVSTMNKEQQMCET PPSESKDILQQVVDHLSPVSTMNKEQQMCET PPSESKDILQQVVDHLSPVSTMNKEQQMCET PPSESKDILQQVVDHLSPVSTMNKEQQMCET PPSESKDILQQVVDHLSPVSTMNKEQQMCET PPSESKDILQQVVDHLSPVSTMNKEQQMCET PPSESKDILQQVVDHLSPVSTMNKEQQMCET PSESKDILQQVVDHLSPVSTMNKEQQMCET PSESKDILQQVVDHLSPVSTMNKEQQMCET PSESKDILQQVVDHLSPVSTMNKEQQMCET PSESKDILQQVDHAVAIVHPSECTYPOLEVY COPARAVINEPSECTYPOLEVY LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS LGGSGTTLSVSEYKKKYREHVLQLHAVKI RAVYYSS PSESKDILTFINLEPSEAAPEEALGPAE PEPGRARRSDTHTFINRIFRRDEEGRPLTVVV QOPAGIGKTMAKKILYDWAAGKILYQGQV PSAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALIGPEBAA CTDPFEAASGARVLGGLISKALLPTALLLVT RAAAPGRIQGRICSPQCAEVRGFSDKKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPQVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTILRGDAQPHSHLVLTT RRLFGLSAERMDDERHFGCMVSRVKQEA LRWVQGQQQCGVAPEVTEGAKGLEDTEE DGGVPRTAAGGVGTILRGDAQPHSHLVLTT RRLFGLSAERMDDERHFGCMVSRVKQEA LRWVQGQQQGCGVAPEVTEGAKGLEDTEE PEEEEGEEEPNYPLELLYCLYETQEDAFVRQ LCRFPELALQRVRFCRMDVAVLSYCVCCP QQALRLISCRLVAAQEKKKSLGKRLQASLC GG GG SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGN LSHIPQQAFSGLYSLKILMLHNQLGGPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS LWLDDNALTEIPS TIPLGLGAADMCAFPWILILLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVAVVLQESISLPLEIPPDEEVENI WRWCGSEEVAVVLQESISLPCEIPPDEEVENI WRWCGSEEVAVVLQESISLPCEIPPDEEVENI WRWCGSEEVAVVLQESISLPCE			1		ļ		FGSKENTLNIKKALLSGYFMQIARDVDGSGN
HKFSISENNYIRITSEISPLFMQLVPQYYFSNI PPSESKDILQQVVDHLSPVSTMNKEQQMCET PPSESKDILQQVVDHLSPVSTMNKEQQMCET PPSESKDILQQVVDHLSPVSTMNKEQQMCET PPSESKDILQQVVDHLSPVSTMNKEQQMCET PPSESKDILQQVVDHLSPVSTMNKEQQMCET PPSESKDILQQVVDHLSPVSTMNKEQQMCET PPSESKDILQQVVDHLSPVSTMNKEQQMCET PSESKDILQQVVDHLSPVSTMNKEQQMCET PSESKDILQQVVDHLSPVSTMNKEQQMCET PSESKDILQQVVDHLSPVSTMNKEQQMCET PSESKDILQQVDHAVAIVHPSECTYPOLEVY COPARAVINEPSECTYPOLEVY LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS LGGSGTTLSVSEYKKKYREHVLQLHAVKI RAVYYSS PSESKDILTFINLEPSEAAPEEALGPAE PEPGRARRSDTHTFINRIFRRDEEGRPLTVVV QOPAGIGKTMAKKILYDWAAGKILYQGQV PSAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALIGPEBAA CTDPFEAASGARVLGGLISKALLPTALLLVT RAAAPGRIQGRICSPQCAEVRGFSDKKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPQVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTILRGDAQPHSHLVLTT RRLFGLSAERMDDERHFGCMVSRVKQEA LRWVQGQQQCGVAPEVTEGAKGLEDTEE DGGVPRTAAGGVGTILRGDAQPHSHLVLTT RRLFGLSAERMDDERHFGCMVSRVKQEA LRWVQGQQQGCGVAPEVTEGAKGLEDTEE PEEEEGEEEPNYPLELLYCLYETQEDAFVRQ LCRFPELALQRVRFCRMDVAVLSYCVCCP QQALRLISCRLVAAQEKKKSLGKRLQASLC GG GG SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGN LSHIPQQAFSGLYSLKILMLHNQLGGPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS LWLDDNALTEIPS TIPLGLGAADMCAFPWILILLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVVAVLQESISLPLEIPPDEEVENI WRWCGSEEVAVVLQESISLPLEIPPDEEVENI WRWCGSEEVAVVLQESISLPCEIPPDEEVENI WRWCGSEEVAVVLQESISLPCEIPPDEEVENI WRWCGSEEVAVVLQESISLPCE		1	1				YLMLTHKOVAOLHPLSGYSITKKMPEWVLF
PSESSKOILQQVVDHLSPVSTMNKEQQMCET CPETEQRCTLQ 754 2104 A 6055 2 394 YYALHHWPPPDLLCQTTGAIFQMNMYGSCII LMLNVDRYAAIVHPLRLRHLRRPRVARLLC LGVWALILVFAVPAARVHRPSRCRYRDLEVI LCFESSFDELWKGRLLPLVLLAEALGFLLPLA AVVYSS 755 2105 A 6059 3 1795 LGLGSGTLLSVSEYKKKYREHVLQLHARVKI RNARSVKITKRPTKLLIAPESAAPEEALGPAE PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVI QGPAGIGKTMAAKKLIYDWAAGKLYQGQVI FAAFFMPGEGLLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAA CTDFFEAASGARVLGGLISKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSVV LLFITSVLSSPVADGPRLQGDLRNLCRLARR GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGGGCGPVAPEVTEGAKGLEDTEE PEPEEEGEGEPNYPLELLVCLYETQEDAFVRQ, LCRFPELALQRVRFCRMDVAVLSYCVRCCP/ GQALRLISCRLVAAQEKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLVSLLIMLHNNQLGGIPAQA LWELPSLQSLRLIDANLISLVPERSFEGLSSLRI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLIDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQESQRRL	1		1				
CPETEQRCTLQ			Į.		l		
Total	ł	l	1		ł		
LMLINVDRYAAIVHPLRLHRLRRPRVARLLC LGVWALILVFAVPARRVHRPSRCRYRDLEVI LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS 755 2105 A 6059 3 1795 LGLGSGTTLLSVSEYKKYREHVLQLHARVKI RNARSVKITKRFTKLLIAPESAAPEEALGPAE PEPGRARRSDTHTFNRLFRRDEEGRRPLT-VVI QGPAGIGKTMAAKKILYDWAAGKLYQGQVI FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRILFILDGADELPALGGPEAA CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFRDERRAERAYFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV: LLFITSVLSSAPVADGPRLQGDLRNLCRLARR GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLGTEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERRDIERHFGCMVSERVKQEA LRWVQGQQGCPGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCP- GQALRLISCRLVAAQEKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGGPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	754	2104	 	6055	2	204	
LGVWALILVFAVPAARVHRPSRCRYRDLEVH LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS 755 2105 A 6059 3 1795 LGLGSGTILSVSEYKKKYREHVLQLHARVKI RNARSVKITKRPTKLLIAPESAAPEEALGPAE PEPGRARRSDTHTFINRLFRDEEGRRPLTVVVI QOPAGIGKTMAAKKILYDWAAGKLYQGQV FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRILFILDGADELPALGGPEAA CTDPFEAASGARVLGGILSKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTISVV LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLSAERMRDIERHFGCMVSERVKQEA LRWVQGQQGCPGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQQ LCRFPELALQRVRFCRMDVAVLSYCVRCCP GQALRISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLVSLKILMLHNNQLGGIFAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGIGAADMCAFPWLLILLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	134	2104	^	6022	2	394	
CCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS Total			-*-	1	1		T
755 2105 A 6059 3 1795 LGLGSGTLLSVSEYKKKYREHVLQLHARVKI RNARSVKITKRPTKLLIAPESAAPEEALGPAE PEPGRARRSDTHTFNRLFRADEEGRRPLTVVI QGPAGIGKTMAAKKILYDWAAGKLYQGQVI FAFFMPCGELERPGTRSLADLILLQCPDRGA PVPQMLAQPQRLLFILDGADELPALGQPEAA CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRGDLSRTSKTTTSVI LLFITSVLSSAPVADGPRLQGDLSRTSKTTTSVI LLFITSVLSSAPVADGPRLQGDLRGSKVQTLFLSK KELPGVLETVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQQGQGCPGVAPEVTEGAKGLEDTEE PEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPAGQALRLISCRLVAAQEKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTQQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRIDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	ļ		1	ļ	ļ		
755 2105 A 6059 3 1795 LGLGSGTLLSVSEYKKKYREHVLQLHARVKI RNARSVKITKRFTKLLIAPESAAPEEALGPAE PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVI QGPAGIGKTMAAKKILYDWAAGKLYQGQVI FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAA CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSVI LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQCGVAPEVTEGAKGLEDTEE PEEEEGEGEPNYPLELLYCLYETQEDAFVRQVLCRFPELALQRVRFCRMDVAVLSYCVRCCP/GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRJDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 TIPLGLGAADMCAFPWLLILLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI			į	İ			
RNARSVKITKRPTKLLIAPESAAPEEALGPAE PEPGRARRSDTHTFINRLFRRDEEGRRPLTVVI QGPAGIGKTMAAKKILYDWAAGKLYQGQVI FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAA CTDPFEAASGARVLGGLLSKALLPTALLLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV' LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQQGCGPQAPEVTEGAKGLEDTEE PEEGEEGEEPNYPLELLYCLYETQEDAFVRQ, LCRFPELALQRVRFCRMDVAVLSYCVRCCP, GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 TIPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		l	<u> </u>			l	AVVYSS
PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVV QGPAGIGKTMAAKKILYDWAAGKLYQGQVV FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAA CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV' LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQCPGVAPEVTEGAKGLEDTE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 TIPLGLGAADMCAFFWLILLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	755	2105	A	6059	3	1795	LGLGSGTLLSVSEYKKKYREHVLQLHARVKE
PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVV QGPAGIGKTMAAKKILYDWAAGKLYQGQVV FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAA CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV' LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQCPGVAPEVTEGAKGLEDTE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 TIPLGLGAADMCAFFWLILLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	1	ļ	ŀ	ļ		Į.	RNARSVKITKRFTKLLIAPESAAPEEALGPAEE
QGPAGIGKTMAAKKILYDWAAGKLYQGQVI FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRILIFILDGADELPALGGPEAA CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV' LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQCCPGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPHELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	1						PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVL
FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAA CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV' LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEBEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	1	İ		ļ			
PVPQMLAQPQRLLFILDGADELPALGGPEAA CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV' LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQCPGVAPEVTEGAKGLEDTEE PEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCP- GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	ļ		ļ	ĺ			
CTDPFEAASGARVLGGLLSKALLPTALLLVT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTITSV' LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI			ł	ļ	1	1	DADOMI VODOBEKI OTKOPUDELDAL CODE VAD
RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV' LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRILDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	-		i	İ		1	
YFYKFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV' LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI			1				
PFVCWIVCTVLRQQLELGRDLSRTSKTTTSV LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		1	1	1		1	
LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	1	1	1	ļ]	9	f .
GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		1	1		\	1	PFVCWIVCTVLRQQLELGRDLSRTSKTTTSVY
GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	1		1				LLFITSVLSSAPVADGPRLQGDLRNLCRLARE
KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		1		}			1
DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		[}	
RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		1				1	
LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		1		1		1	
PEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		1	ł	1	1	1	
LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		1]	}		1	
GQALRLISCRLVAAQEKKKKSLGKRLQASLC GG 756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMILHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		1	1	1			
756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMILHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		1	1				LCRFPELALQRVRFCRMDVAVLSYCVRCCPA
756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMILHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		1	1	1	1	1	GQALRLISCRLVAAQEKKKKSLGKRLQASLG
756 2106 A 6060 12 436 SGRPTRPAKPTGQGMGRFMLTLVCQGSIMM ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMILHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI							GG
ARDLIMNNLTELQPGLFHHLRFLEELRLSGNI LSHIPGQAFSGLYSLKILMILHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLILLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	756	2106	† A	6060	12	436	1
LSHIPGQAFSGLYSLKILMILHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLILLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	1	-100	1	*****		1 .55	
LWELPSLQSLRLDANLISLVPERSFEGLSSLRI LWLDDNALTEIPS 757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI)	1	}	}	1	I SHIDGUARSCI VSI KII NII UNINOI CCIDACA
LWLDDNALTEIPS LWLDDNALTEIPS 1757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI	1	1		1		1	
757 2107 A 6063 54 419 ITPLGLGAADMCAFPWLLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		1	1.	1			
WRWCGSEEVVAVLQESISLPLEIPPDEEVENI		<u> </u>		<u> </u>	ļ. <u>.</u>		
	757	2107	A	6063	54	419	
				1	1	·	WRWCGSEEVVAVLQESISLPLEIPPDEEVENII
	l		1		1		WSSHKSLATVVPGKEGHPATIMVTNPHYQG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QILTMLLRSLQQPSASWPRDCSSSCSW
758	2108	Α	6066	125	438	IGISCPATIFVPMFSHSLIGIGEEYQLPYYNMV PSDPSYEDMREVVCVKRLRPIVSNRWNSDEC LRAVLKLMSECWAHNPASRLTALRIKKTLAK MVESQDVKI
759	2109	A	6072	3	650	PGRRFRPAALEERAMEKLREKVPFQNRGKGT LSSIIPNNSDTRKATETTSLSSKPEYVNPDFRW SKDPSSKSGNLLETSEVGWTSNPEELDPIRLA LLGKSGLSCQVGSATSHPVSCQEPIDEDQRISP KDKSTAGREFSGQVSHQTTSENQCTPIPSSTV HSSVADMQNMPAAVHALLTQPSLSAAPFAQ RYLGTLPSTGSTTLPQCHAGNATVW
760	2110	A	6077	3	730	PLRLTLMEEVLLLGLKDREGYTSFWNDCISSG LRGCMLIELPLRGRLQLEACGMRRKSLLTRK VICKSDAPTGDVLLDEALKHVKETQPPETVQ NWIELLSGETWNPLKLHYQLRNVRERLAKNL VEKGVLTTEKQNFLLFDMTTHPLTNNNIKQR LIKKVQEAVLDKWVNDPHRMDRRLLALIYL AHASDVLENAFAPLLDEQYDLATKRVRQLLD LDPEVECLKANTNEVLWAVVAAFTK
761	2111	A	6078	833	390	IVSFHLSGFKKFVRPFSFLSVHGLQVDEYHSV HQKLSADMADHSNLIRSLLVGAEDARLMRD MKTMKSRYMELYDLNRDLLNGYKIRWNNH TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT ACRDAIRSNNINTLFKIMRVGTASS
762	2112	A	6079		2686	KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS HSDLWSSSSSLESSSFPLPKQYLDVSSQTDISG SFGINSNNQLAEKVRLRLRYEEAKRIANLKI QLAKLDSEAWPGVLDSERDRLILINEKEELLK EMRFISPRKWTQGEVEQLEMARKRLEKDLQ AARDTQSKALTERLKLNSKRNQLVRELEEAT RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ SKVEFLLLEGATGFRPSGCTITHEDEVAKTQ KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA TLCELSLGNSAQERYRLEEPGTEGKQLGQAV NTAQGCGLKVACVSAAVSDESVAGDSGVYE ASVQRLGASEAAAFDSDESEAVGATRIQIALK YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG AQISLAEVCRSGERSTRWYNLLSYKYLKKQS RELKPVGVMAPASGPASTDAVSALLEQTAVE LEKRQEGRSSTQTLEDSWRYEETSENEAVAE EEEEEVEEEGEEDVFTEKASPDMDGYPALK VDKETNTETPAPSPTVVRPKDRRVGTPSQGPF LRGSTIIRSKTFSPGPQSQYVCRLNRSDSDSST LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS VLKELKEQLEQAKSHGEKELPQWLREDERFR LLLRMLEKRMDRAEHMGELQTDKMMRAAA KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR MNIPALSADDV
763	2113	A	6082	3	1558	PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE VLENLTQGKMCLVPGKTRKLLFKFVAKTED VGKKIEITSVDLALGNETGRCVVLNWQGGGG DAASSQEALQAARSFKRRPKLPDNEVHWGSII IQASTMIISRVPNISVHLLHEPPALTNEMYCLV

الله النبيا النبيا النبير البياد و سند سد

NO. of No. of body body household coation of corresponding sequence wence where the coation of t	CEA ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide sequence USSN 09496 per uence unce correspondit procession of the process of the proc	SEQ ID					i	
uence Sequence			1100				
Sequence		1					
uence Pida anino acid a famino acid residue of peptide poptide peptide peptide peptide sequence Pida		-					M=Methionine N=Asparagine P=Proline
amino acid residue of peptide requence propriet sequence provide sequence	-	donce)				
residue of peptide sequence V=Tyrosine, X=Ushanown, ==Siep codon, Poposible nucleotide delicion, \	dence			314			
peptide sequence sequ							
		ļ		ļ		Sequence	
VIVQSHEKTQRDWKLTAGLKRGQDANLTQK HVVILITGELCDESYPALLTDIPVGDLHPGEQ LEKM_VYRCGTVGSRMFLVYVSYLNTTVEE REIVCKCHKDETVITETVFFPDVAVKYSYKRTK ERLERVYADPFLIMTDLLSASPWALTTVSSE LELAPANTTVDQLSSYWANLTQTSSE LLAPANTTVDQLSSYWANLTQTSSE LLAPANTTVDQLSSYWANLTQTSSE LLAPANTTVDQLSSYWANLTQTSSE LLAPANTTVQDLSSYWANLTQTSSE LLAPANTTVQDLSSYWANLTQTSSE LLAPANTTVQDLSSYWANLTQSASE LAPANTTVQDLSSYWANLTQSASE LAPANTTVQDLSSYWANLTQSASE LAPANTTVQDLSSYWANLTQSASE LAPANTTVQDLSSYWANLTQSASE LAPANTTVQDLSSYWANLTQSASE LAPANTTVQDLSSYWANLTQSASE LAPANTTVQDLSSYWANLTQSASE LAPANTTVQDLSSYWANLTQSASE LAPANTTVQDLSSYSYANLTQSASE LAPANTTVQDLSSYNGSYSYANLTQSASE LAPANTTVQDLSSYNGSYSYANLTQSASE LAPANTTVQDLSSYNGSYSYSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS				Ì			
THYTLHOTELCDESYPALLIDIPVGDLHPGEG LEKM, VYRGGTVGSKRHLVYSYS, UNITVEE EERIVCKCHKDETVTETVFFPDVAVKFVSTKF EHLERYVADPFLLMTDLLSASPWALTVISSE LHAPSMTTVDQLESQVDNVILOTGESASECF CLQCPLGMIEGGVATGHVISWKRTSAMENI PHTTVTTLPHVIVENPLLBVNADLPSFGRVERS LPVKYHLQNKTDLVQDVEISVEPSDAFMFSG LKQRIRILPGTEQBML*NNYPLMAGYQQLPS LNNRLRFPHSTDLVQDVEISVEPSDAFMFSG LKQRIRILPGTEQBML*NNYPLMAGYQQLPS LNNRLRFPHSTDLVQDVEISVEPSDAFMFSG LKQRIRILPGTEQBML*NNYPLMAGYQQLPS LNNRLRFPHSTDLVQDVEISVEPSDAFMFSG LKQRIRILPGTEQBML*NNYPLMAGYQQLPS LNNRLRFPHSTDLVQCATKATOPATHSV VALVASSSYPJAMFYQDLAEVKSEHIPDVEPA AFADLANSNAGAAVGRKAGPRSPFSAPAPAP PPAPAPPTLONNHQESPGWECCRPTLRENN ALMFNNELMADVHFVVGPPGATRTVPAHKY VALVASSSYPJAMFYQDLAEVKSEHIPDVEPA AFILLKYMYSDEIDLEADTVLATLYAAKKYI VALVAKSVPJAMFYAGRALVVALVAIPTMTLEE EEPELTQRCWEVIDAQAEMALRSEGFCEIDR GTLEIUTREALNTKEAVVFEAVLINWAEAEC KRQGLPITPRIKRRVLGRALVVALVAIPTMTLEE EPELTQRCWEVIDAQAEMALRSEGFCEIDR GTLEIUTREALNTKEAVVFEAVLINWAEAEC KRQGLPITPRIKRRVLGRALVVARFATINKPALD PPITRKRGLAPQRCHRFQSSAYRSNOWRYRG RCDSIGFADVREVFTAMFYATA RCDSIGFADVRAVARFATISHISTURVATARTITISHISTURVATA			L	ļ	sequence		
LEKMLYVKGGTVGSRMFLVYSVLINTTVEE ELLRAYYADIPFLIMTDILSASPWALITIVSE ELLRAYYADIPFLIMTDILSASPWALITIVSE ELLRAYYADIPFLIMTDILSASPWALITIVSE ELLRAYYADIPFLIMTDILSASPWALITIVSE ELLPASMITTVDQLESQVDNVLIGTGESASECP CLQCSI,GINEGGVATGHYIISWKRTSAMEN PITTVTTI,PHYVTENPLHYNADIPSPGRVRES LEVKYHILONKTDLVQDUEISVEPSDAFMFSG LEVKYHILONKTDLVQDUEISVEPSDAFMFSG LEVKYHILONKTDLVQDUEISVEPSDAFMFSG LEVKYHILONKTDLVQDUEISVEPSDAFMFSG LEVKYHILONKTDLVQDUEISVEPSDAFMFSG LEVKYHILONKTDLVQDUEISVEPSDAFMFSG LEVKYHILONKTDLVQDUEISVEPSDAFMFSG LEVKYHILONKTDLVQDUEISVEPSDAFMFSG LEVKYHILONKTDLVQDUEISVEPSDAFMFSG LEVKYHILONKTDLVQDUEISVEPSDAFMFSG LEVKYHILONKTDLVQDUEISVEPSDAFMFSG LEVKYHILONKTDLVQDUEISVEPSDAFMFSGAFMFSFAFMFSGAFMFSGAFMFSGAFMFSGAFMFSFAFMFSGAFMFSFAFMFSGAFMFSFAFMFSGAFMFSFAFMFSGAFMFSGAFMFSGAFMFSFAFMFSGAFMFSFAFMFSGAFMFSGAFMFSGAFMFSFAFMFSGAFMFSGAFMFSGAFMFSFAFMFSGAFMF					i		
		İ	l	1		ļ	
				İ			
LHLAPSMITUDQLESQUANULIQTGESASE		ĺ	ĺ	ĺ			
CLQCPSLGNIEGÓVATGHYIISWKRTSAMEN							
PHITTUTILPHVIVADILPHOVADILPSGORIVES		ļ	ł	Į			
1		l		1	1]	CLQCPSLGNIEGGVATGHYIISWKRTSAMENI
LKORILRIPGTEQEMLYNFYPIMAGYQQLPS		ŀ	ļ		}	j	
1422 AAADLANSNAGAAVGRKAGPRSPSARPAP PPDAPAPPTLGNNHQESPGWRCCRPTLRENN ALMFNNELMADVHFVVGPGATRIVPAHKY VIAVGSSVPYAMFYGDLAEVKSEBHPDVEPA AFILILKYMYSDEDLEADTVLATLYAAKKYI VPALAKACVNFLETSLEAKNACVLLSGSLF EPPELTQRCWEVIDAQAMALRSEGFCEIDR EPPELTQRCWEVIDAQAMALRSEGFCEIDR CTLLIKYMYSDEDLEADTVLATLYAAKKYI VPALAKACVNFLETSLEAKNACVLLSGSLF EPPELTQRCWEVIDAQAMALRSEGFCEIDR EPPELTQRCWEVIDAQAMALRSEGFCEIDR CTLLIKYMYSDEDLEADTVLATLYAAKKYI VPALAKACVNFLETSLEAKNACVLLSGSLF EPPELTQRCWEVIDAQAMALRSEGFCEIDR EPPELTQRCWEVIDAQAMALRSEGFCEIDR CROSSIGRAVSRYRG CROSSIGRAVSR CR			ļ			ì	LPVKYHLQNKTDLVQDVEISVEPSDAFMFSG
DDTSIAAA		İ	İ	į.			LKQIRLRILPGTEQEMLYNFYPLMAGYQQLPS
1422			ļ	l			LNINLLRFPNFTNQLLRRFIPTSIFVKPQGRLM
PPPAAPPTLGNNHOESPGWRCCRPTLRERN ALMFNNELMADVHFVVGPPGATRTVPAHKY			ì	1		}	DDTSIAAA
ALMFNNELMADVÍEVVGPPGATRTVPAHKY VLAVGSSVYJAMPIGDLAPVKSEHIPDVSPA AFILLK YMYSDEIDLEADTVLATILYAAKKYI VPALAKACVNFLETSLEAKNACVLLSQSRLF EEPELTQRCWEVIDAGMALTSEGFECIDR QTLEILVTREALNTKEAVVFEAVLNWAEAEC KRQGLPITPRNKRHVLORAT-VLVREPTMTLEE FANGAAQSDILTLEETHSIFLWYTATNKPRLD FPLTKRKGLAPQRCHRFQSSS YRSNQWRYPG RCDSIGFAVDRKVFLAGGLI YGSSSGKAEYSV KIELKRLGVVLAQNI-TKFMSDGSSNTFPVWF EHPVQVEQDTFYTASAVLDGSELSYFGQEGM TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA 765 2115 A 6099 I 1150 SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHFDVNVWEASGNRSGGVCDDCQHN TEGQCYGCKVGFFYBAPCKFCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYSGCRCDCAGSCU PTIGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKLSRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYSCRCDCAGSCU PTIGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKLABHGKTHVEVVNKIK KVLKSTKLLKIFRGKRTLYPESWTDRGCTCPL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK FSLGRKVMDILKRECK THE A 6103 2 384 MTAAATATVLKEGVLEKRSGGLIQLWKRK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK FSLGRKVMDILKRECK TVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVLESTGRHINFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 766 2116 A 6106 1 542 SGSSHASDGSGFQELRICSEDQTPLLAGMCSLP MARYYIIKYADQKALYTRDGQLLVGPPVAD NCCAEKICTLPNRGLDRTKVPFILGIQGGSRC LACVETEEGPSLQLEDVNIELLYKGGEATRF TFFQSSSGSAFRLEAAAWFGWFLCGFAEPQQ PVQLTKESEFSARIKFYFEQSW FFFQVSSGSAFRLEAAAWFGWFLCGFAEPQQ PVQLTKESEFSARIKFYFEQSW FFFMVNPHIVGVKTKGUFLSSFEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYRAPSSSSSSSS SSSSSSSSSNSNTSDSTKSQHNSTSNIVNGAY ETSDSEMBMEAEHTPNGVLGSMSTRIVNGAY ETSDSEMBMEAEHTPNGVLGSMSTRIVNGAY ETSDSEMBMEAEHTPNGVLGSMSTRIVNGAY ETSDSEMBMEAEHTPNGVLGSMSTRIVNGAY ETSDSEMBMEAEHTPNGVLGSMSTRIVNGAY	764	2114	Α	6093	1	1422	AAADLANSNAGAAVGRKAGPRSPPSAPAPAP
ALMFNNELMADVÍEVVGPPGATRTVPAHKY VLAVGSSVYJAMPIGDLAPVKSEHIPDVSPA AFILLK YMYSDEIDLEADTVLATILYAAKKYI VPALAKACVNFLETSLEAKNACVLLSQSRLF EEPELTQRCWEVIDAGMALTSEGFECIDR QTLEILVTREALNTKEAVVFEAVLNWAEAEC KRQGLPITPRNKRHVLORAT-VLVREPTMTLEE FANGAAQSDILTLEETHSIFLWYTATNKPRLD FPLTKRKGLAPQRCHRFQSSS YRSNQWRYPG RCDSIGFAVDRKVFLAGGLI YGSSSGKAEYSV KIELKRLGVVLAQNI-TKFMSDGSSNTFPVWF EHPVQVEQDTFYTASAVLDGSELSYFGQEGM TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA 765 2115 A 6099 I 1150 SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHFDVNVWEASGNRSGGVCDDCQHN TEGQCYGCKVGFFYBAPCKFCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYSGCRCDCAGSCU PTIGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKLSRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYSCRCDCAGSCU PTIGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKLABHGKTHVEVVNKIK KVLKSTKLLKIFRGKRTLYPESWTDRGCTCPL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK FSLGRKVMDILKRECK THE A 6103 2 384 MTAAATATVLKEGVLEKRSGGLIQLWKRK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK FSLGRKVMDILKRECK TVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVLESTGRHINFTLVTEGGGEDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 766 2116 A 6106 1 542 SGSSHASDGSGFQELRICSEDQTPLLAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPFILGIQGGSRC LACVETEEGPSLQLEDVNIELLYKGGEATRF TFFQSSSGSAFRLEAAAWFGWFLCGFAEPQQ PVQLTKESEFSARIKFYFEQSW FFFQSSGSSAFRLEAAAWFGWFLCGFAEPQQ PVQLTKESEFSARIKFYFEQSW FFFMVNPHIVGVLKTKQULSSGEATRFTGCVSHIGALLAF TTFSVSSSMHRVARCAAPHVHILLANFYLLF PPMVNPHIVGVLKTKQULSSGEATRFTGCVSHIGALLAF TFTSVSSSMHRVARCAAPHVHILLANFYLLF PPMVNPHIVGVLKTKQULSSGEATRFTNNGGALSFFEKGCVNRE TSDSEMBMEAEHTPNGVLGSMSTRIVNGAY ETSDSEMBMEAEHTPNGVLGSMSTRIVNGAY ETSDSEMBMEAEHTPNGVLGSMSTRIVNGAY ETSDSEMBMEAEHTPNGVLGSMSTRIVNGAY	1	1	l	ł	1]	
VLAVGSSVFYAMFYGDLAEVKSEHIPDVEPA							,
AFILILKYMYSDEIDLEADTVLATLYAAKKYI VPALAKACVNFLETSLEAKNACVLLSQSRLF EEPELTQRCWEVIDAQAEMALRSEGFCEIDR QTLEIIVTREALNTKEAVVFEAVLNWAEAEC KRQGLPITPRNKRHVLGRAVLVREPTMTLEE FANGAAQSDILTLEETHSIFLWYTATNKPRLD FPLTKRKGLAPQRCHRFQSSAYRNQWRYRG RCDSIQFAVDRRVFLAGLGLYGSSSGKAEYSV KIELKRLGVVLAQNLTKFMSDGSNTFPVWF EHPVQVGDTFYTASAVLDGSELSYFGGEGM TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA 65 2115 A 6099 1 1150 SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF RPVKAPGTFHMVHGKCMCKHNTAGSHCOH CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHPDVNVWEASGRRSGGVCDDCQHN TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTPDSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALHASKKTPSSNTDSTKSCCCKPGTTGNAKA FCGMKYSYVLKIKLSAHDKGTHVEVNVKIK KVLKSTKLKJFRGKRTLYPESWTDRGCTCPL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK 766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVSTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQOAJGTVRARQSLGTGTL VS SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPHICIQGGSRC LACVETEEGFSLQLEDVNIEELYKGGESATF TFFQSSSGSAFRLEAAAWPGWFLGPAEPQQ PVQLTKESEFPSARTKFYFEGSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTTSVISSVMHRVARCAAPHVHILLANFYLLF PMVNPHIYGVKTKQHSJGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAFSSSSSSS SSSSSSSSSNSVNYSESNSTDSTKSQCHNESTSNQ			J	}			
PALAKACVNFLETSLEAKNACVLLSQSRLF EEPELTQRCWEVIDAQAEMALRSEGFCEIDR QTLEIIVTREALNTKEAVVFEAVLNWAEAEC KRQGLPITPRNKRIVLGRALYLVRIPTIMTLEE FANGAAQSDILTLEETISILWYTATNKPRLD FPLTKRKGILAPQRCHRFQSSAYRSNQWBYRG RCDSIQFAVDRRVFIAGLGLYGSSSGKAEYSV KIELKRLGVVLAQNLTIKFMSDGSSNTFPVWF EHPVQVEQDTFYTASAVLDGSELSYFGQEGM TEVQCQKVAFQPGCSSDSTNGTGVQGGQIPE LIFYA 765 2115 A 6099 1 1150 SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHFDVNVWBASGMRSGGVCDDCQHN TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSVVLKIKILSAHDKGTHVEVNVKIK KVLKSTKLKJFRGKRILYPESWTDRGCTCPL NPGLFYLVAGHEDIRTGKLIVNMKSFYVQHWK PSLGRKVMDILKRECK 766 2116 A 6103 2 384 MTAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 1 542 SGSSHASDGSGFQELICSEDQTPLLAGMCSLP MARYYIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEGFGSLLQLEVNIEELYKGGEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFGGSW 768 2118 A 6109 3 292 FILQAVLQLSSQEATKFYFGGSW PSVLTKSESPSARTKFFGTCVSHIGALLAF YTTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIFIYGVKKGNGLSISPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSNSVNYSESNSTDSTKSCHHSSTSNQ ETSDSEMBEMEAEHYPNGVLGSMSTEKSCHNESTSNQ ETSDSEMBEMEAEHYPNGVLGSMSTEKSCHNESTSNQ ETSDSEMBEMEAEHYPNGVLGSMSTEKSCHNESTSNQ ETSDSEMBEMEAEHYPNGVLGSMSTEKSCQNRE	1	1	ļ	1			
BEPELTQRCWEVIDAQAEMALRSEGFCEIDR QTLEIIVTREALNTKEA VVFEAVLNWAEAEC RQGLPIIPRNKRHVLGRAL YLVRIPTIMTLEE FANGAAQSDILTLEETHSIFLWYTATNKPRLD FPLTKRKGLAPQRCHRFQSSAYRSNQWRYRG RCDSIQFAVDRRVFIAGLGLYGSSSGKAEYSV KIELKRLGVVLAQNLTKFMSDGSSNTFPVWF EHPVQVQQDTFYTJSAVLDGSSELSYFGQEGM TEVQCGKVARQFQCSSDSTNGTGVQGGQIPE LFYA 6099 1	1						
OŢILEIIVĪREALNĪKEA VVFRAVLNIWAEABC			1				
				1			· · ·
FANGAAQSDILTLEETHSIFLWYTATNKPRLD FPLTKRKGLAPQRCHRFQSSAYRSNQWRYRG RCDSIQFAVDRAVILAGQLITKMSDGSSNTFPVWF EHPVQVEQITPYTASAVLDGSELSYFGQEGM TEVQCGKVAFQFQCSSDSTNGTGVQGQIPE LIFYA SGFTHVAIVDFIVKGSCFCNVHADQCIPVHGF RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPMECRTCKCNG CAPLYNDRPWEAADGKTGAPMECRTCKCNG CAPLYNDRPWEAADGKTGAPMECRTCKCNG GRRCDRCMVGYWGFGDYGCFPCDCAGSCD PTGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKLSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK FSLGRKVMDILKRECK MTAAATATVLKEGVLEERSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHYFTLVTEGGGEDFRCPLEDFGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTIRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPTLGIQGGSCD PVQLTKESEFSARIKFVFEQSW CACKEKICTLPNRGLDRTKVPTLGIQGGGSCD CACKEKICTLPNRGLACH CACKEKICTLPNRGLDRTKTPTLORGD C	}		į	1	}	Į	
FFLTKRKGLAPQRCHRFQSSAYRSNQWRYRG RCDSIQFAVDRRVFIAGLGLYGSSSGKAEYSV KELLKRLGVVLAQNLTKFMSDGSSNTFPVWF EHPVQVEQDTFYTASAVLDGSELSYFGQEGM TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA 1150 SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHFDVNVWEASGNRSGGVCDDCQHN TEQQYCQRCKPGFYRDLRRFFSAPDACKPCS CHPVGSAVLFANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALHSGKCECKEQTLGNAKA FCGMKYSYVLKIKLISAHDKGTHVEVNVKIK KVJLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVSTGRHIYFTLVTEGGGEDFRCPLEDFGW NAQITLGLVKFKNQQIQTVRARQSLGTGTL VS SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNGLDRTKVPFLGQGGSRC LACVETEGPSLQLEDVNIEELYRGGEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARIKFYFEQSSW TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARIKFYFEQSSW TFFQSSSGSARILEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARIKFYFEQSSW TFFQSSSGSARRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARIKFYFEQSSW TFFQSSSGSARRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARIKFYFEQSSW TFFQSSSGSARRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARIKFYFEQSSW TFFQSSSGSARIKFAAAAWPGWFLCGPAEPQQ PVQLTKESEPSARIKFYFEQSSW TFFQSSSGSARIKFAAAAWPGWFLCGPAEPQQ PVQLTKESEPSARIKFYFEQSSW TFFQSSSGSFRIKGAAAWPGWFLCGPAEPQQ PVQLTKESEPSARIKFYFEQSSW TFFQSSSGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS		j		i	l		
RCDSIOFADDREVFILAGIGLYGSSSGKAEYSV KIELKRLGVVLAQNLTKFMSDGSSNTFPVWF EHPVQVEQDTPYTASAVLDGSELSYFQQEGM TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA				ŀ	l		
KIELKRI, GVLA, QNLTKFMSDGSSNTFPVWF EHPVQVEQDTFYTASAVLDGSELSYFGQEGM TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHFDVNVWEASGNRSGGVCDDCQHN TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKLSAHDKGTHVEVNVKIK KVLKSTKLKIFGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK TOTAL			i	}			
HPVQVEQDTFYTASAVLDGSELSYFGQEGM				1			
TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA	ł	l	1	}	ł	ł	
1150 SGFHYAITOFIVKGSCFCNVHADQCIPVHGF				1			
765 2115 A 6099 1 1150 SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF				İ			
RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHFDVNVWEASGNRSGGVCDDCQHN TEGQYQCRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEZYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK 766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS GSSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSQNHSSYPAPSSSSSSS SSSSSSSPSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	765	2115	 	6000	1	1150	
CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHFDVNVWEASGNRSGGVCDDCQHN TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKLISAHDKGTHVEVNVKIK KVLKSTKLKIFGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKYMDILKRECK 766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQAIQTVRARQSLGTGTL VS 767 2117 A 6106 1 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHYHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSSSNSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	103	2113	Ι^ .	0099	1	1130	1
HADTCHFDVNVWEASGNRSGGVCDDCQHN TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKLSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK 766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 I 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPITYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSQNHISKYPAPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	[1	I	ĺ		
TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK 766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 1 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSQNHSKYPAPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	İ			1		·	
CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTI.GNAKA FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK TO 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS TO SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSAARTKFYFEQSW TO 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPITYGVKTKQIRDSLGSIPEKGCVNRE TO 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY			l	1			
GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKLSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK 766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 1 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEGGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAFSSSSSSS SSSSSSSSSSSSYSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	}		ł	l	l	i	
PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK 766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 I 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPII'YGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSSSSTSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	1				1		
WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK 766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 1 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSQHHSKTPAPSSSSSS SSSSSSSSSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY				1	I		
FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK 766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 1 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEBATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	1				1	1	
KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK 766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 I 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YPPVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 I 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY			1	1	1	1	
NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK 766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 1 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSGARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSSS SSSSSSSSSSSSSSNSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	ľ	-	1	1 .	ľ	l	·
PSLGRKVMDILKRECK 766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 1 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPITYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSSSSNNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY			1		I	l	
766 2116 A 6103 2 384 MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 1 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSSSSSSSSSSSSSSSSSSS]			1			
CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 l 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 l 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY			L	<u> </u>		<u> </u>	
CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 l 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 l 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSSSSSSSNSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	766	2116	Α	6103	2	384	1
CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS 767 2117 A 6106 l 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 l 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSSSSSSSNSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	1		[ſ		[
767 2117 A 6106 I 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 I 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSSSSNSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	1	1	ì	1	1		
767 2117 A 6106 I 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSS SSSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY						1	
767 2117 A 6106 1 542 SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY		1	1	ł	1		1 ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' '
MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	767	2117	A	6106	1	542	
NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	1		ļ		-	1	
LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY		ļ	ſ	i			NCCAEKICTI PNRGI DRTKVPIEI GIOGGEPO
TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW 768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	1	1		1	1	1	
768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	1	1	1		1	1	
768 2118 A 6109 3 292 FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE 769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY		1	1	i	Į.		
769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSSSSSSSSSSSSSSSSSSS	760	2112	 	(100	12	202	
769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSSSSSSSSSSSSSSSSSSS	/08	2118	A	6109	13	292	
769 2119 A 6110 1 711 RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSSSSSSSSSSSSSSSSSSS			1			1	
SSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY		ļ	<u> </u>	<u> </u>		L	
SSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY	769	2119	A	6110	1	711	RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS
ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY		1	Į.				SSSSSSPSSVNYSESNSTDSTKSQHHSSTSNQ
KHEDI OTDESSMODRHPRROI CCGNOA ATE					Į.		ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY
		1	}		1	1	KHEDLQTDESSMDDRHPRRQLCGGNQAATE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first		Q=Ghutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
1	ĺ	ľ		residue of	sequence	/=possible nucleotide deletion, \=possible
		1		peptide sequence		nucleotide insertion
	 -	ļ	 	sequence		RIILFGRELQALSEQLGREYGKNLAHTEMLQD
						AFSLLAYSDPWSCPVGQQLDPIQREPVCAAL
ſ	[[-	ĺ		NSAILESQNLPKQPPLMLALGQASECLRLMA
			İ			RAGLGSCSFARVDDYLH
770	2120	Ā	6125	2	570	YFGLNLHVQHLGNNVFLLQTLFGAVILLANC
1 110	2120	A	0123	4	370	VAPWALKYMNRRASQMLLMFLLAICLLAIIF
				1		VPQEMQMLREVLATLGLGASALANTLAFAH
1					1	GNEVIPTIIRARAMGINATFANIAGALAPLMM
						ILSVYSPPLPWIIYGVFPFISGFAFLLLPETRNK
1	ŀ	ł			i	PLFDTIQDEKNERKDPREPKQEDPRVEVTQF
771	2121	A	6126	909	353	RSFVLDTASAICNYNAHYKNHPKYWCRGYF
''1	2121	^	0120	100	, , , , ,	RDYCNIIAFSPNSTNHVALRDTGNQLIVTMSC
	}]]	1		LTKEDTGWYWCGIQRDFARDDMDFTELIVT
						DDKGTLANDFWSGKDLSGNKTRSCKAPKVV
	1	l		1		RKADRSRTSILIICILITGLGIISVISHLTKRRS
1				i	}	QRNRRVGNTLKPFSRVLTPKEMAPTEQM
772	2122	A	6148	7	810	FVLGILALSHTISPFMNKFFPASFPNRQYOLLF
1112	2122	^	0140	l ′	010	TOGSGENKEEINYEFDTKDLVCLGLSSIVGV
		l				WYLLRKHWIANNLFGLAFSLNGVELLHLNN
i i		1		([VSTGCILLGGLFIYDVFWVFGTNVMVTVAKS
i			1	İ		FEAPIKLVFPQDLLEKGLEANNFAMLGLGDV
						VIPGIFIALLLRFDISLKKNTHTYFYTSFAAYIF
1		j	ļ	J]	GLGLTIFIMHIFKHAQPALLYLVPACIGFPVLV
				ł		ALAKGEVTEMFSYEESNPKDPAAVTESKEGT
			1			EASASKGLEKKEK
773	2123	A	6161	3	1088	COPMLYTRKNHPKLLLRRTESVAEKMLTNW
1113	2125	A	0101	3	1000	FTFLLYKFLKESAGEPLFMLYCAIKHQMEKG
1		ļ			1	PIDAITGEARYSLSEDKLIRHLIDYKTLTLNCV
	'					NPENENAPEVPVKGLDCDTGTQAKEKLLDA
			1			AYKGVPYSQRPKAADMDLEWRQGRMARIIL
ł		l		l		ODEDVTTKIDNDWKRLNTLAHYOVTDGSSV
		ŀ				ALVPKQTSAYNISNSSTFTKSLSRYESMLRTA
1						SSPDSLRSRTPMITPDLESGTKLWHLVKNHDH
i	· ·					LDQREGDRGSKMVSEIYLTRLLATKGTLQKF
1	1	[1	1	VDDLFETIFSTAHRGSALPLAIKYMFDFLDEQ
		ı				ADKHQIHDADVRHTWKSNCLPLRFWVNVIK
1			ļ	1		NPQFVFDIHKNSITDACLSVV
774	2124	A	6163	860	125	KTAVKKRNLNPVFNETLRYSVPQAELQGRVL
' ' ' '	2124	^	0103	300	123	SLSVWHRESLGRNIFLGEVEVPLDTWDWGSE
1			1		1	PTWLPLOPRVPPSPDDLPSRGLLALSLKYVPA
1				i		GSEGAGLPPSGELHFWVKEARDLLPLRAGSL
1	1	l	1	1	1	DTYVQCFVLPDDSRASRQRTRVVRRSLSPVF
1		[1	1	1	NHTMVYDGFGPADLRQACAELSLWDHGALA
1						NRQLGGTRLSLGTGSSYGLQVPWMDSTPEEK
1	ļ			1	l	QLWQALLEQPCEWVDGLLPLRTNLAPRT
775	2125	A	6191	2	392	ARGIGSLGRDHSGSGGGTGMAGAWVRKAAD
1113	2123	^	0171	١٠	332	YVRSKDFRDYLMSTHFWGPVANWGLPIAAIT
			1	l	1	DMK/KSPEIISRRMTFAL*CYSLTFVRFAHYVQ
1	1		1			
1	1	1	1	[VPWNWLMLGCHTAVDFDQLISSMPCISHGMT ASASAL
776	0126	 	(217	 , 	927	ASASAL FRGYWGVREAFTDASWSGGLGPGKPGMKJT
776	2126	Α	6217	1	827	t ·
1	1	1		1	}	ROKHAKKHLGFFRNNFGVREPYQILLDGTFC
		1	1		1	QAALRGRIQLREQLPRYLMGETQLCTTRCVL
	1		I	1		KELETLGKDLYGAKLIAQKCQVRNCPHFKNA
1		1	1			VSGSECLLSMVEEGNPHHYFVATQDQNLSVK
1	1	1	1	I	1	VKKKPGVPLMFIIQNTMVLDKPSPKTIAFVKA
	}	1	1	1	1	VESG\RLSQCMRKKVSNISKRNRV**KTLNRG
]	!	1		RRKKRKKISGPNPLSCLKKKKKAPDTQSSASE
		1	1	i	1	KKRKRKRIRNRSNPKVLSEKQNAEGE

Г	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
1	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
۱	nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
	eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
١	seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
١	uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1					amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
-					residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
					peptide		/=possible nucleotide deletion, \=possible
L					sequence		nucleotide insertion
-	777	2127	A	6236	1038	1402	YYQISSLPSIVGNGIFLWLLICIFLAKQGGSRL*
- [1						FQPFGRPRGGGHLRSGVLGQPGQHGETP/SFF
-				:			YNSKISPALWGPPVIPSALGGEAGKSL*PRRQ
Į							RFQRGGIAPLPSRVRGRAKLFLKKK
	778	2128	A	6237	422	913	ASFFHHHRGAFLLLLAIPGS*GQDQSLIHWSN
-							AVSNAD\LLDLK\N*LDH\LEEKMPL\EVKVVP
- [PQVL\SEPN*RSGGCFSAPSFEVPPWTGEVKP/
1							SPQRDGGALG\QGPLGIPSDSILALLKKQT*RA
1							LLNWPLGSLRRSSCFGGQDGQDLKPRSGLGC
-		0100					NSFRYRR
-	779	2129	A	6249	420	36	ARAPSPSFSVRDVELSDPARERGEMPVAVGP
	.						YGQSQPSCFDRVKMGFVMGCAVGMAAGAL
ļ							FGTFSCLSSILVSSSG/SGMRGRELMGGIGKTM
1							MQSGGTFGTFMAIGMGIRC*PWLPTTSVPSH
ŀ	700	2120		6262	415	1200	QSQPMY
	780	2130	A.	6263	415	1380	RIMRMCDRGIQMLITTVGAFAAFSLMTIAVG
							TDYWLYSRGVCRTKSTSDNETSRKNEEVMT
ļ							HSGLWRTCCLEGAFRGVCKKIDHFPEDADYE
							QDTAEYLLRAVRASSVFPILSVTLLFFGGLCV
ŀ							AASEFHRSRHNVILSAGIFFVSAGLSNIIGIIVYI
							S\ANAGRTPGQR\DSKKSYSYGWSF/YFSGAFS
							FIIGR/IIC*GVGLPWHIYIEKHQQLRAKSHSEF
		'					LKKSTFARLPPYRYRFRRRSSSRSTEPRSRDLS
							PISKGFHTIPSTDISMFTLSRDPSKITMGTLLNS
1	ł						DRDHAFLQFHNSTPKEFKESLHNNPANRRTT
-	781	2131	Α —	6274	832	318	PV RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH
	701	2131	А	0274	632	316	LQVKQKQLACLCTWQARDPDCPPSTKVVL/L
-				· I			VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS
1							QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD
-							LEFLGDLKGCSELKNFOELITOSALVHPKADV
							WWYCGRPLLGTLPSN
ŀ	782	2132	A	6281	1324	393	WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG
١		-10-		0201	132.		EREDWGIGSA*SVGAVSKVPSARF*RTYPS\E
J							DEEEVTHQKSSSSDSNSEEHRKKKTSRSRNK
							KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD
1							DDKKRVKAKKKKKKKKKKKKKKKKKKKKKKK
							ESSDSSCKDSEEDLSEATWMEQPNVADTMDL
]							IGPEAPIIHTSQDEKPLKYGHALLPGEGAAMA
							EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM
							SGSRHRRMEAVRLRKENQIYSADEKRALASF
							NQEERRKRESKILASFREMVHKKTKGKDDK
ı	783	2133	A	6305	201	1032	WDDYPQGALRRREAAEGLHFLGPPGRVRGQ
-					,		LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP
							AMARPPVPGSVVVPNWHES/RRGQGVPGLHS
							AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV
						}	SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ
							TTVVFWPAKLQASSRVVMFRFEFWDCGESA
-]	LKKFDHMLLACMENTDAFLFLFSFTDRASFE
							DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT
							DVPERDLTAFRQAWELPLLRVKSVPGRRLG
ſ	784	2134	Α	6308	86	96	GSSPDPASLITMKNQDKKNGAAKOSNPKSSP
- 1							GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA
							PRKPEGAQARTAQSGALRDVSEELSROLEDIL
							STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR
-[-	TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR
						r ·	1
							QSDEVGDRDHRRPQEKKKAKGLGKEITLLM
							QSDEVGDRDHRRPQEKKKAKGLGKEITLLM QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ KQMKLLQKKQSQLVQEKDHLRGEHSKAVLA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RSKLESLCRELQRHNRSLKEEGVQRAREEEE KRKEVTSHFQVTLNDIQLQMEQHNERNSKLR QENMELAERLKKLIEQYELREEHIDKVFKHK
	·					DLQQQLVDAKLQQAQEMLKEAEERHQREKD FLLKEAVESQRMCELMKQQETHLKQQLALY TEKFEEFQNTLSKSSEVFTTFKQEMEKMTKKI KKLEKETTMYRSRWESSNKALLEMAEEKTV RDKELEGLQVKIQRLEKLCRALQT/GAQ*PVR GQRWGSHRTSAVRIFS
785	2135	A	6319	1493	889	SPQGPLLRSVSPVSAGASSVTPGGAQPGVTTT PPSLVAVAPAPGSAAGPAAGWQ*HAGCR/WT KLPWSWGMRPMKIFFSEEYRSISTRISHDAL* EKCTQPAKPLSMIR\TGSSVSPG/PLVKWNWT RREFRNSGTRVVSSCCGMSCMYSFLGHCSV/S QDLPLVHVDVGWQPPLGPTVGLRPGLLPLHD TTPCQKLVVDDLDWA
786	2136	A	6320	551	135	RWLPVAECDSSCVGCTGEGPGNCKECISGYA REHGQCADVDECSLAEKTCVRKNENCYNTP GSYVCVCPDGFEET/RRCLCAAGRG*SHRRRK PDTAALPRRPVMCRTYPLNYSEGCPVENVAL RMPSPAVDSGGERLPAL
787	2137	A	6330	1693	227	DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG SGILGLAYVMANTGVFGFSFLLLTVALLASYS VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTCL PLIEFLQSL*NSL*AVTSYEDLGLFAFGLPGKL VVAGTIIIQNIGAMSSYLLIIKTELPAAIAEFLT GDYSRYWYLDGQTLLIIICVGIVFPLALLPKIG FLGYTSSLSFFFMMFFALVVIIKKWSIPCPLTL NYVEKGFQISNVTDDCKPKLFHFSKESAYALP TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE VTCHRIKDKVESELLKG***IP*SHDVVVMT\V KLCILFAVLL\TVPLIHFPARKAVTMMFFSNFP FSWIRHFLITLALNIIIVLLAIYVPDIRNVFGVV GASTSTCLIFIFFGLFYLKLSREDFLSWKKLGV GCFC/LLSFKTSILRNSLSVYILLPASRKSIYFKI
788	2138	A	6351		6622	PRSLCFSLWAEAAVLADGGLRRRRILLRGTM SASFVPNGASLEDCHCNLFCLADLTGIKWKK YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV LG/WRRDQRPERRE\L*IFWGGEDP\VLLTLF TMTYQKKKMECGRMDFPMNAVLCFSKAVH NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ AFKMSDSATKKLIGEWKQFYPISCCLKEMSE EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV DRVWQECNMNRAQNKRKYSASSGGLCEEAT AAKVASWDFVEATQRTNCSCLRHKNLKSRN AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL VISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE MANSPQPPPLSPHPCDVVDEGVTKTPSTPQS QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ YQEAVEPTVYVGTAVNLEEDEANIAWKYYK FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV TSVTELMVQCKKPLKVSDELVQQYQIKNQCL SAIASDAEQEPKIDPYAFVEGDEEFLFPDKKD

	700 TO 1					
(SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	Ì			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						RQNSEREAGKKHKVEDGTSSVTVLSHEEDA
						MSLFSPSIKQDAPRPTSHARPPSTSLIYDSDLA
						VSYTDLDNLFNSDEDELTPGSKRSANGSDDK
						ASCKESKTGNLDPLSCISTADLHKMYPTPPSL
						EQHIMGFSPMNMNNKEYGSMDTTPGGTVLE
						GNSSSIGAQFKIEVDEGFCSPKPSEIKDFSYVY
						KPENCQILVGCSMFAPLKTLPSQYLPLIKLPEE
						CIYRQSWTVGKLELLSSGPSMPFIKEGDGSNM
						DQEYGTAYTPQTHTSCGMPPSSAPPSNSGAGI
			1			LPSPSTPRFPTPRTPRTPRTPRGAGGPASAOGS
						VKYENSDLYSPASTPSTCRPLNSVEPATVPSIP
						EAHSLYVNLILSESVMNLFKDCNSDSCCICVC
1						NMNIKGADVGVYIPDPTQEAQYRCTCGFSAV
į l						MNRKFGNNSGLFFEDELDIIGRNTDCGKEAE
						KRFEALRATSAEHVNGGLKESEKLSDDLILLL
į l						QDQCTNLFSPFGAADQDPFPKSGVISNWVRV
						EERDCCNDCYLALEHGROFMDNMSGGKVDE
						1
				- !		ALVKSSCLHPWSKRNDVSMQCSQDILRMLLS
						LQPVLQDAIQKKRTVRPWGVQGPLTWQQFH
i i		ľ				KMAGRGSYGTDESPEPLPIPTFLLGYDYDYLV
1						LSPFALPYWERLMLEPYGSQRDIAYVVLCPE
						NEALLNGAKSFFRDLTAIYESCRLGQHRPVSR
						LLTDGIMRVGSTASKKLSEKLVAEWFSQAAD
						GNNEAFSKLKLYAQVCRYDLGPYLASLPLDS
						SLLSQPNLVAPTSQSLITPPQMTNTGNANTPS
						ATLASAASSTMTVTSGVAISTSVATANSTLTT
						ASTSSSSSNLNSGVSSNKLPSFPPFGSMNSNA
						AGSMSTQANTVQSGQLGGQQTSALQTAGISG
						ESSSLPTQPHPDVSESTMDRDKVGIPTDGDSH
1						AVTYPPAIVVYIIDPFTYENTDESTNSSSVWTL
1 1						GLLRCFLEMVQTLPPHIKSTVSVQIIPCQYLLQ
1						PVKHEDREIYPOHLKSLAFSAFTOCRRPLPTS
]						TNVKTLTGFGPGLAMETALRSPDRPECIRLYA
						PPFILAPVKDKQTELGETFGEAGQKYNVLFV
						GYCLSHDQRWILASCTDLYGELLETCIINIDVP
[]						NRARRKKSSARKFGLQKLWEWCLGLVQMSS
						LPWRVVIGRLGRIGHGELKDWSCLLSRRNLQ
j						SLSKRLKDMCRMCGISAADSPSILSACLVAM
[]						EPQGSFVIMPDSVSTGSVFGRSTTLNMQTSQL
						NTPQDTSCTHILVFPTSASVQVASATYTTENL
					}	, -
						DLAFNPNNDGADGMGIFDLLDTGDDLDPDII
1						NILPASPTGSPVHSPGSHYPHGGDAGKGQSTD
[RLLSTEPHEEVPNILQQPLALGYFVSTAKAGP
ļ l						LPDWFWSACPQAQYQCPLFLKASLHLHVPSV
; l						QSDELLHSKHSHPLDSNQTSDVLRFVLEQYN
j l						ALSWLTCDPATQDRRSCLPIHFVVLNQLYNFI
L						MNML
789	2139	Α	6359	1	2002	TGTLTEDGLDVMGVVPLKGQAFLPLVPEPRR
						LPVGPLLRALATCHALSRLQDTPVGDPMDLK
						MVESTGWVLEEEPAADSAFGTQVLAVMRPP
1		!				LWEPQLQAMEEPPVPVSVLHRFPFSSALQRM
1						SVVVAWPGATQPEAYVKGSPELVAGLCNPET
						VPTDFAQMLQSYTAAGYRVVALASKPLPSVP
1 1						SLEAAQQLTRDTVEGDLSLLGLLVMRNLLKP
1						QTTPVIQALRRTRIRAVMVTGDNLQTAVTVA
j l						RGCGMVAPQEHLIIVHATHPERGQPASLEFLP
, l						MESPTAVNGVKDPDQAASYTVEPDPRSRHLA
1						I THE TATE OF THE PARTY OF THE PERSON OF THE
						I SCOTECHIVE HEDRI I DEVI VOCTUBADAGAD
·				_		LSGPTFGIIVKHFPKLLPKVLVQGTVFARMAP
				-		LSGPTFGIIVKHFPKLLPKVLVQGTVFARMAP EQKTELVCELQKLQYCVGMCGDGANDCGAL KAADVGISLSQAEASVVSPFTSSMASIECVPM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VIREGRCSLDTSFSVFKYMALYSLTQFISVLIL YTINTNLGDLQFLAIDLVITTTVAVLMSRTGP ALVLGRVRPPGALLSVPVLSSLLLQMVLVTG
			,			VQLGGYFLTLAQPWFVPLNRTVAAPDNLPNY ENTVVFSLSSFQYLILAAAVSKGAPFR\RPLTN NVPFLLASAL*SSVLVVLVLSPGLLHGPLALR NITDTGFKLLLVGLVTLNFVGGLHAGERARP VPPRLPAPPPAQAG\SKKRFKQLERELAEQPW PPLPAGPLR
790	2140	A	6380	76	1059	SSAGSARKLQVMALAARLWRLLPFRRGAAP GSRLPAGTSGSRGHCGPCRFRGFEVMGNPGT FKRGLLLSALSYLGFETYQVISQAAVVHATA KVEEILEQADYLYESGETEKLYQLLTQYKESE DAELLWRLARASRDVAQLSRTSEEEKKLLVY EALEYAKRA/L/EKNESSFASHKWYAICLSDV GDYEGIKAKIANAYIIKEHFEKAIELNPKDATS IHLMGIWCYTFAEMPWYQRRIA*NACLQLPP *FPPYEKALGYFHRAEQVDPNFYSKNLLLLG KTYLKLHNKKLAAFWLMKAKDYPAHTEED KQIQTEAAQLLTSFSEKN
791	2141		6434	3	1460	IALLIVDGLAWDDQGGLALLHISPSKLIL*QDS SGMS/YVMVRCTITRAFFKSLLCHICQYSIGPQ *VT\CPGQDACKE*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEAEALARVYSSSISNGLSNLN NETSGTYANGSVIDLPKSEGYYNVVSGQPSP DQSGLDMT\GIKQIKQEPIYDLTSVPNLFTY\SS FNN\GQLAPGIT\MTEIDRIAQNIIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK
792	2142	A	6440	92	781	SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL
793	2143	A	6446	3201	152	PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD EPAMKRPRTAGPAVHLKAMQLSWTSLALVG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
		,				IDSHGKLSVLRLSPSMGHPLEVGLALRHLLFL LEYCMVTGYDWWDILLHVQPSMVQSLVEKL HEEYTRQTAALQQVLSTRILAMKASLCKLSP CTVTRVCDYHTKLFLIAISSTLKSLLRPHFLNT PDKSPGDRLTEICTKITDVDIDKVMINLKTEEF VLDMNTLQALQQLLQWVGDFVLYLLASLPN QPCPTSEPCPTSEPSPTSEPSPTSEPSSP*SLC\G SLLRPGHSFLRDGTSLGMLRELMVVIRIWGLL KPSCLPVYTATSDTQDSMSLLFRLLTKLWICC RDEGPASEPDEALVDECCLLPSQLLIPSLDWL PASDGLVSRLQPKQPLRLQFGRAPTLPGSAAT LQLDGLARAPGQPKIDHLRRLHLGACPTEEC KACTRCGCVTMLKSPNRTTAVKQWEQRWIK NC/LVRWALVAGAPQLPLSPAAPQLLLSYPSA APEPGCCKSHRSPWTLLGAVNLSPPCRAVEG RGPDACVTSRASEEAPAFVQLGPQSTHHSPRT PRSLDHLHPEDRP
794	2144	A	6490	418	585	NGDKADLENESCRAQVLMPVVPALWEAEGG
795	2145	A	6499	395	1027	GSIEPRDLRLQ*AVITPL\TPAWVTQ KLLWLPPHSEQKRSPLYHPQGPSGTTPSAP\FS SHSPPPSLLQA\PSIAAFLRTHGHISASGPLRMP FPH\H*NAFLLVFPGQRSQLTS\PSHYLCREVFP DHHHHLCRLSLESSPLFHHRVLFCVPKQNVN STRAQIFCLFVHIVGCRCINTFPLHLFRLHLWL HFLQIPLCKKNKSVKLGKTVVGRGCQSAAGS DTRVRAAVGAPGLPVEPLV
796	2146	A	6503	68	936	HSALLTHSSFCVFTLCQDFFTYSSMSEEVTYA DLQFQNSSEMEKIPEIGKFGEKAPPAPSHVWR PAALFLTLLCLLLIGLGVLASMFHVTLKIEM KKMNKLQNISEELQRNISLQLMSNMNISNKIR NLSTTLQTIATKLCRELYSKEQEHKCKPCPRR WIWHKDSCYFLSDDVQTWQESKMACAAQN ASLLKINNKNALEFIKSQSRSYDYWLGLSPEE DS/YSWYESG*YNQVPSAWVIRNAPDLNNMY CGYINRLYVQYYHCTYKQRMICEKMANPVQ LGSTYFREA
797	2147		6507	i	881	PGSTHASARSQVPRSAGEAAPHSRRPPGLLPH APRAASAQLEERMRDPHPGMTLQEGDCRGS QTVSLTMGTADSDEMAPEAPQHTHIDVHIHQ ESALAKLLLTCCSALRPRATQARGSSRLLVAS WVMQIVLGILSAVLGGFFYIRDYTLLVTSGA AIWTGAVAVLAGAAAFIYEKRGGTYWALLR TLLALAAFSTAIAALKLWNEDFRYGYSYYNS ACRISSSSDWNTPAPTQSPEEVRRLHLCTSFM DMLKALFRTLQAMLLGVWILLLASLTPLWL /SL/RGECSQPKG*VPKKRDQKEMLEVSGI*PG STHASARSQVPRSAGEAAPHSRRPPGLLPHAP RAASAQLEERMRDPHPGMTLQEGDCRGSQT VSLTMGTADSDEMAPEAPQHTHIDVHIHQES ALAKLLLTCCSALRPRATQARGSSRLLVASW VMQIVLGILSAVLGGFFYIRDYTLLVTSGAAI WTGAVAVLAGAAAFIYEKRGGTYWALLRTL LALAAFSTAIAALKLWNEDFRYGYSYYNSAC RISSSSDWNTPAPTQSPEEVRRLHLCTSFMDM LKALFRTLQAMLLGVWILLLLASLTPLWLYC WRMFPTKGVSP
798	2148	A	6528	912	2287	VPNYLPSVSSAIGGEVPQRYVWRFCIGLHSAP RFLVAFAYWNHYLSCTSPCSCYRPLCRLNFG LNVVENLALLVLTYVSSSEDF/TWVPG*GRSG

			1 000		1 20 10 11 1	L.A
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	l	in	nucleotide	location	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	1	USSN	location	corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence		l	914	ng to first		T=Threonine, V=Valine, W=Tryptophan,
		l		amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			residue of	sequence	/=possible nucleotide deletion, \=possible
	İ	1	ŀ	peptide		
	L		ļ	sequence		nucleotide insertion
			i			EVFPEGTGLPLPHSDLPTSWCGHSLQCGSQSS
		1				FPPAIHENAFIVFIASSLGHMLLTCILWRLTKK
		1				HTVSQE\DGLSLAGAPRQPRRKSRTSVLRIRV
	ĺ	1	1			MVRWELSSNGNPGRGVLGLGLGLGNKLRVV
]	j	ļ	GQNLGL*HCVWVVWETGE*KRWRLQMGIE*
		İ		ļ		GVASRRQ*VRNSVRGLVCHNSSAPPMYMGFF
		-	1			SPTVFGGGVGG*LHVTFILHPPEVEAAGIPLLL
		l	i			GPSLPQRQGREHIVVILAAPACAPFHDR*WEP
		l				REIRPSP*ELGLRGEPTLSYPASCRVIRQPIP*D
		İ				RKSYSWKQRLFIINFISFFSALAVYFRHNMYC
		1		ļ		EAGVYTIFAILEYTVVLTNMAFHMTAWWDF
		<u> </u>				GNKELLITSQPEEKRF
799	2149	Α	6529	1	874	FFFFQRINFIEHSGSVSLLALACDLGWCEDWS
		1				CCLVQGGGDLVDVVQTNHGEDEAGGDTDSV
						DEARCKESQQEAQENLREDLCLESFAKDKIL
		1				QIIEGSEREHEETRTKQAALDGEPLGGGQLTA
		Ī				VHLHPSKEQQGQEGGERQRGARTHHWRGW
				_		EKGRRVRLRPPSGKLRADQPVRKLGGPTPS/T
						ELPGLQPHAPTPHTA/PATPTYSPAPDTPNPPV
		•	1			RWKCPLPVEPRTRQLCRERTRKACPPKPRPPL
						GLPGDPTGPVTHHAPPVSPTGASGQERRAEP
						GAVSYAHASATK
800	2150	Α	6544	2	662	SAQRWAAVAGRWGCRLLALLLLVPGPGGAS
			1	İ		EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG
		}	1	1.		GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF
					!	TASKNGTYKFCFSNE\FSTFTHKTVYFDFQVG
	İ					E\THLCFLVR/DRVSALTQMESACVSIHEALKS
			1			VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV
	İ	Í	Ĺ	J	·	GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS
801	2151	A	6556	1	1319	TPCMECIKGEGLREPQNLSGSQREPQTEGSM
				<u> </u>		DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF
				1		KRIFLKRMPSIRESLKERGVDMARLGPEWSQP
ļ		1	1			MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP
			1		'	PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY
						HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL
1				İ		SQDIITVGGITVTQMFGEVTEMPALPFMLAEF
ł		1	1	1	1	DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED
			1			VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE
		1	1	1		GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE
	1		}	1		DGCLALVDTGASYISGSTSSIEKLMEALGAKE
l	1		İ			KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT
1			1			SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL
	<u> </u>	<u></u>		<u> </u>		\ALGATF\IRKFYTEFDRGNNPHGFALAR
802	2152	Α	6567	13	6147	MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL
1		1	1	1		LAVVVLLALPVAWGQCNAPEW\LPFARPTNL
!		1	j	1		TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS
			1	1		VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG
1		1	1	1		IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW
[1	1			DNETPICDRIPCGLPPTITNGDFISTNRENFHY
İ		1	1	1	1	GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND
1		1	1			DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
	1	1				NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA
l		1				LNKWEPELPSCSRVCQPPPDVLHAERTQRDK
l		1	1		1	DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG
		1		+		DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV
1		1				NLQLGAKVDFVCDEGFQLKGSSASYCVLAG
1		1		1		MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP
1		1			1	LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR
				1		CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL
1	1	1	1	1		FAKLKTQTNASDFPIGTSLKYECRPEYYGRPF
1						

000.00	GEA TO	N 4	1 000	Des die Ac d	Dandier 3 - 3	Lamino paid company (A=Ala-15: C. C. a.
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of nucl-	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
		1	USSN			
eotide	seq-	i	1	location	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi		M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		Ì		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
		1			1	SITCLDNLVWSSPKDVCKRKSCKTPPDPVNG
			!			MVHVITDIQVGSRINYSCTTGHRLIGHSSAECI
					İ	LSGNAAHWSTKPPICQRIPCGLPPTIANGDFIS
					į	TNRENFHYGSVVTYRCNPGSGGRKVFELVGE
						PSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
		}				ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGP
			1			RRVKCQALNKWEPELPSCSRVCQPPPDVLHA
		İ			•	ERTQRDKDNFSPGQEVFYSCEPGYDLRGAAS
			Ì	1		MRCTPQGDWSPAAPTCEVKSCDDFMGQLLN
			ĺ	i		GRVLFPVNLQLGAKVDFVCDEGFOLKGSSAS
						YCVLAGMESLWNSSVPVCEQIFCPSPPVIPNG
						RHTGKPLEVFPFGKAVNYTCDPHPDRGTSFD
		l	ł		}	LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC
						QAPDHFLFAKLKTQTNASDFPIGTSLKYECRP
		ļ				EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP
			1	1		PDPVNGMVHVITDIQVGSRINYSCTTGHRLIG
	l	1	ļ		ļ	HSSAECILSGNTAHWSTKPPICQRIPCGLPPTI
						ANGDFISTNRENFHYGSVVTYRCNLGSRGRK
'		1	ì		}	VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN
		1				l
i				1		KCTPPNVENGILVSDNRSLFSLNEVVEFRCQP
						GFVMKGPRRVKCQALNKWEPELPSCSRVCQ
						PPPEILHGEHTPSHQDNFSPGQEVFYSCEPGY
	į				i	DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF
	ŀ				ŀ	LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL
		ļ]	ļ	KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN
		l	ļ		Į	PPAILNGRHTGTPSGDIPYGKEISYTCDPHPDR
			1			GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE
			1		1	LSVRAGHCKTPEQFPFASPTIPINDFEFPVGTS
]	LNYECRPGYFGKMFSISCLENLVWSSVEDNC
					1	RRKSCGPPPEPFNGMVHINTDTQFGSTVNYSC
		j			ļ	NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII
	Ì	ļ		Ì		SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH
l		ŀ	- .	·		TGPDGEQLFELVGERSIYCTSKDDQVGVWSS
		i				PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI
		1				IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH
						CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY
						SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV
			İ			KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC
:		[!		DEGFRLKGRSASHCVLAGMKALWNSSVPVC
		 			j	EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT
		1				CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS
	ļ		1	1		SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL
		l '	1		1	YLPGMTISYTCDPGYLLVGKGFIFCTDOGIWS
	1	1	1			\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \
		1	1	[QLDHYCKEVNCSFPLFMNGISKELEMKKVYH
		l	1			YGDYVTLKCEDGYTLEGSPWSQCQADDRWD
		1	1			PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI
		ł				ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP
		L		L		RTLQTNEENSRVLP
803	2153	A	6574	2	3233	HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL
	1	1	1	[LLLPPLLLLRG\SHAGNLTVAVVLPLANTSY
			Ī	ļ		PWSWA\RVGPAVELALAQVKARPDLLPGWT
	1	1		1		VRTVLGSSENALGVCSDTAAPLAAVDLKWE
		1	1	1		HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL
	1	1	l			TAGAPALGFGVKDEYALTTRAGPSYAKLGDF
	1		!	l		VAALHRRLGWERQALMLYAYRPGDEEHCFF
,		I	1	1	,	LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT
	1	1	l .			
			i .			RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA
						RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW
						RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						DGLLLYIQAVTETLAHGGTVTDGENITQRMW NRSFQGVTGYLKIDSSGDRETDFSLWDMDPE NGAFRVVLNYNGTSQELVAVSGRKLNWPLG YPPPDIPKCGFDNEDPACNQDHLSTLEVLALV GSLSLLGILIVSFFIYRKMQLEKELASELWRVR WEDVEPSSLERHLRSAGSRLTLSGRGSNYGSL LTTEGQFQVFAKTAYYKGNLVAVKRVNRKR IELTRKVLFELKHMRDVQNEHLTRFVGACTD PPNICILTEYCPRGSLQDILENESITLDWMFRY SLTNDIVKGMLFLHNGAICSHGNLKSSNCVV DGRFVLKITDYGLESFRDLDPEQGHTVYAKK LWTAPELLRMASPPVRGSQAGDVYSFGIILQE IALRSGVFHVEGLDLSPKEIIERVTRGEQPPFR PSLALQSHLEELGLLMQRCWAEDPQERPPFQ QIRLTLRKFNRENSSNILDNLLSRMEQYANNL EELVEERTQAYLEEKRKAEALLYQILPHSVAE QLKRGETVQAEAFDSVTIYFSDIVGFTALSAE STPMQVVTLLNDLYTCFDAVIDNFDVYKVET IGDAYMVVSGLPVRNGRLHACEVARMALAL LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGEAL\KI HLSS\ETKAVL\EEFGGFELERGDVEMKGKG
804	2154	A	6585	2	3837	DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV MSERVSGLAGSIYREFERLIVRYDEEVVKELIP LVVAVLENLDSVFAQDQEHQVELELLRDDNE QLITQYEREKALRKHAEEKFIEFEDSQEQEKK DLQTRVESLESQTRQLELKAKNYADQISILEE REAELKKEYNALHQRHTEMIHNYMEHLERT KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP AGDGLLTPDAQKGGETPGSEQWKFQELSQPR SHTSLKDELSDVSQGGSKATTPASTANSDVA TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL FEELSSAGSGLIGDVDEGADLLGMGREVENLI LENTQLLETKNALNIVKNDLIAKVDELTCEK DVLQGELEAVKQAKLKLEEKNRELEEELRKA RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE MARVLMERNQYKERLMELQEAVRWTEMIR ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK KPEPPVNILKYNAPTSHVTPSVKKRSSTLSQLP GDKSKAFDFLSEETEASLASRREQKREQYRQ VKAHVQKEDGRVQAFGWSLPQKYKQVTNG QGENKMKNLPVPVYLRPLDEKDTSMKLWCA VGVNLSGGKTRDGGSVVGASVFYKDVAGLD TEGSKQRSASQSSLDKLDQELKEQQKELKNQ EELSSLVWICTSTHSATKVLIIDAVQPGNILDS FTVCNSHVLCIASVPGARETDYPAGEDLSESG QVDKASLCGSMTSNSSAETDSLLGGITVVGC SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEEATEATEGNAGSAEDTVDIS QTGVYTEHVFTDPLGVQIPEDLSPVYQSSND SDAYKDQISVLPNEQDLVREEAQKMSSLLPT MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD SILSIVHVKGIVLVALADGTLAIFHRGVDGQW DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ DVDIEPYVSKMLGTGKLGFSFVRITALMVSC

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		İ	1	peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
		[1	sequence		nucleotide insertion
			 	sequence		NRLWVGTGNGVIISIPLTETVILHQGRLLGLR
	ļ		ļ			ANKTSGVPGNRPGSVIRVYGDENSDKVTPGT
				•		FIPYCSMAHAQLCFHGHRDAVKFFVAVPGQV
	ļ]	ļ			ISPQSSSSGTDLTGDKGRGHLHRSLVVRRP
805	2155	A	6605	469	2602	FGRLLWGTAFKSWKMKAPIPHLILLYATFTQ
						SLKVVTKRGSADGCTDWSIDIKKYQVLVGEP
						VRIKCALFYGYIRTNYSLAQSAGLSLMWYKS
			ļ			SGPGDFEEPIAFDGSRMSKEEDSIWFRPTLLQ
		l	1			DSGLYACVIRNSTYCMKVSISLTVGENDTGL
	ļ	1				CYNSKMKYFEKAELSKSKEISCRDIEDFLLPT
			ļ			REPEILWYKECRTKTWRPSIVFKRDTLLIREV
		1			[REDDIGNYTCELKYGGFVVRRTTELTVTAPL
	1		1		1	TDKPPKLLYPMESKLTIQETQLGDSANLTCRA
	1	[FFGYSGDVSPLIYWMKGEKFIEDLDENRVWE
	1		Į			SDI/KILKEHLGEQEVSISLIVDSVEEGDLGNYS
			1			CYVENGNGRRHASVLLHKRELMYTVELAGG
			1			LGAILLLLVCLVTIYKCYKIEIMLFYRNHFGA EELDGDNKDYDAYLSYTKVDPDOWNOETGE
	}					EERFALEILPDMLEKHYGYKLFIPDRDLIPTGT
	,]	1	İ		YIEDVARCVDQSKRLIIVMTPNYVVRRGWSIF
	1	ì	ļ			ELETRLRNMLVTGEIKVILIECSELRGIMNYQE
				•		VEALKHTIKLLTVIKWHGPKCNKLNSKFWKR
]					LQYEMPFKRIEPITHEQALDVSEQGPFGELQT
	}		ŀ	ļ	†	VSAISMAAATSTALATAHPDLRSTFHNTYHS
						QMRQKHYYRSYEYDVPPTGTLPLTSIGNQHT
		ſ	1			YCNIPMTLINGQRPQTKSSREQNPDEAHTNSA
						ILPLLPRETSISSVIW
806	2156	Α	6614	3	1584	NSARGGVGVRGARAMATVQEKAAALNLSAL
						HSPAHRPPGFSVAQKPFGATYVWSSIINTLQT
		i	1	Ì	ļ	QVEVKKRRHRLKRHNDCFVGSEAVDVIFSHL
			İ		ļ	IQNKYFGDVDIPRAKVVRVCQALMDYKVFE
						AVPTKVFGKDKKPTFEDSSCSLYRFTTIPNQD
			1			SQLGKENKLYSPARYADALFKSSDIRSASLED
						LWENLSLKPANSPHVNISTTLSPQVINEVWQE ETIGRLLQLVDLPLLDSLLKQQEAVPKIPQPK
		1	1			ROSTMVNSSNYLDRGILKAYSDSQEDEWLSA
				1		AIDCLEYLPDQMVVEISRSFPEQPDRTDLVKE
	1					LLFDAIGRYYSSREPLLNHLSDVHNGIAELLV
		ł	İ	ł	,	
	1	1				NGKTEIALEATOLLLKLLDFONREEFRRILLYF
					ļ	NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV
						MAVAANPSEFKLQKESDNRMVVKRIFSKAIV
						MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT
						MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKK\LLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALHHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMLLDLSHNALE
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMILDLSHNALE TLELGARALG\SLRTLLLQGNALRDLPPYTFA
807	2157	A	6615	4198	2094	MAVAANPSEFKLQKESDNRMVVKRIFSKAIV DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD LSGNQLRSILASPLGFYTALRHLDLSTNEISFL QPGAFQALTHLEHLSLAHNRLAMATALSAG GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMLLDLSHNALE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LDLSSNPGLEVATGALGGLEASLEVLALQGN GLMVLQVDLPCFICLKRLNLAENRLSHLPAW TQAVSLEVLDLRNNSFSLLPGSAMGGLETSLR RLYLQGNPLSCCGNGWLAAQLHQGRVDVDA TQDLICRFSSQEEVSLSHVRPEDCEKGGLKNI NLIIILTFILVSAILLTTLAACCCVRRQKFNQQ YKA
808	2158	A	6619	153	1852	FKALSQYIYTNTHLEREAAFEVAILLRRMEEG ARHRNNTEKKHPGGGESDASPEAGSGGGGV ALKKEIGLVSACGIIVGNIIGSGIFVSPKGVLEN AGSVGLALIVWIVTGFITVVGALCYAELGVNI PKSGGDYFYVKDIFGGLAGFLRLWIAVLVIYP TNQAVIALTFSNYVLQPLFPTCFPPESGLRLLA AICLLLLTWVNCSSVRWATRVQDIFTAGKLL ALALIIIMGIVQICKGEYFWLEPKNAFENFQEP DIGLVALAFLQGSFAYGGWNFLNYVTEELV DPYKNLYPRAIFISIPLVTFVYVFANV/ALYVT AMSPQELLASNAVAVTFGEKLLGVMAWIM PISVALSTFGGVNGSLFTSSRLFFAGAREGHLP SVLAMIHVKRCTPIPALLFTCISTLLMLVTSD MYTLINYVGFINYLFYGVTVAGQIVLRWKKP DIPRPIKINLLFPIIYLLFWAFLLVFSLWSEPVV CGIGLAIMLTGVPVYFLGVYWQHKPKCFSDFI ELLTLVSQKMCVVVYPEVERGSGTEEANED MEEQQQPMYQPTPTKDKDVAGQPQP
809	2159	A	6621	1041	223	ODSRKMLPSTSVNSLVQGNGVLNSRDAARH TAGAKRYKYLRRLFRFRQMDFEFAAWQMLY LFTSPQRVYRNFHYRKQTKDQWARDDPAFL VLLSIWLCVSTIGFGFVLDMGFFETIKLLLWV VLIDCVGVGLLIATLMWFISNKYLVKRQSRD YDVEWGYAFDVHLNAFYPLLVILHFIQLFFIN HVILTDTFIGYLVGNTLWLVAVGYYIYVTFL GYSVGLLFFSVALPFLKNTVILLYPFAPLILLYG LSLALGWNFTHTLCSFYKYRVK
810	2160	A	6623	160	822	SPASGHCRLNGAAVAMFGCLVAGRLVQTAA QQVAEDKFVFDLPDYESINHVVVFMLGTIPFP EGMGGSVYFSYPDSNGMPVWQLLGFVTNGK PSAIFKISGLKSGEGSQHPFGAMNIVRTPSVAQ IGISVELLDSMAQQTPVGNAAVSSVDSFTQFT QKMLDNFYNFASSFAVSQ/VPDDTQ/RPSEMF IPANVVLKWYENFQRRTSTEPSLLENIIWIKIN F
811	2161	A	6627	18	3367	LEGSLNTERAKYYLTITMPHFTVTKVEDPEEG AAASISQEPSLADIKARIQDSDEPDLSQNSITG EHSQLLDDGHKKARNAYLNNSNYEEGDEYF DKNLALFEEEMDTRPKVSSLLNRMANYTNLT QGAKEHEEAENITEGKKKPTKTPQMGTFMG VYLPCLQNIFGVILFLRLTWVVGTAGVLQAF AIVLICCCCTMLTAISMSAIATNGVVPAGGSY FMISRALGPEFGGAVGLCFYLGTTFAAAMYIL GAIEIFLVYIVPRAAIFHSDDALKESAAMLNN MRVYGTAFLVLMVLVVFIGVRYVNKFASLFL ACVIVSILAIYAGAIKSSFAPPHFPVCMLGNRT LSSRHIDVCSKTKEINNMTVPSKLWGFFCNSS QFFNATCDEYFVHNNVTSIQGIPGLASGIITEN LWSNYLPKGEIIEKPSAKSSDVLGSLNHEYVL VDITTSFTLLVGIFFPSVTGIMAGSNRSGDLKD AQKSIPIGTILAILTTSFVYLSNVVLFGACIEGV VLRDKFGDAVKGNLVVGTLSWPSPWVIVIGS FFSTCGAGLQSLTGAPRLLQAIAKDNIIPFLRV

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uciicc		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
- 1		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
				,		
				sequence	 	nucleotide insertion
1				ì		FGHSKANGEPTWALLLTAAIAELGILIASLDL
						VAPILSMFFLMCYLFVNLACALQTLLRTPNW
I						RPRFRYYHWALSFMGMSICLALMFISSWYYA
]		IVAMVIAGMIYKYIEYQGAEKEWGDGIRGLS
1						LSAARFALLRLEEGPPHTKNWRPQLLVLLKL
ľ		Ì				DEDLHVKHPRLLTFASQLKAGKGLTIVGSVIV
ļ		ĺ	1		ĺ	GNFLENYGEALAAEQTIKHLMEAEKVKGFCQ
1				ļ	ļ	LVVAAKLREGISHLIQSCGLGGMKHNTVVM
		1			1	GWPNGWRQSEDARAWKTFIGTVRVTTAAHL
						ALLVAKNISFFPSNVEQFSEGNIDVWWIVHDG
				1	ł	GMLMLLPFLLK\QHKVWRKCSIRFF\TVAQLE
		1	1			DNSIQMKKDLATFLYHLRIEAEVEVVEMHDS
		ļ	J]	j	DISAYTYERTLMMEQRSQMLRHMRLSKTER
{		1			1	DREAQLVKDRNSMLRLTSIGSDEDEETETYQ
					Í	EKVHMTWTKDKYMASRGQKAKSMEGFQDL
i						LNMRPDQSNVRRMHTAVKLNEVIVNKSHEA
		İ		ļ		KLVLLNMPGPPRNPEGDENYMEFLEVLTEGL
		İ	l	1	Ī	ERVLLVRGGGSEVITTYS
812	2162	Α	6628	66	640	AVCTMSEMAELSELYEESSDLQMDVMPGEG
						DLPQMEVGSGSRELSLRPSRSGAQQLEEEGP
						MEEEEAQPMAAPEGKRSLANGPNAGEQPGQ
		1				VAGADFESEDEGEEFDDWEDDYDYPEEEQLS
i		l			1	GAGYRVSAALEEADKMFLRTREPALDGGFO
. 1						MHYEKTPFDQLAFIEELF\SLMVVNRLTEELG
		1		İ		CDELIDRE
813	2163	A	6630	708	1355	AKMGAYKYIQELWRKKQSDVMRFLLRVRC
013	2103	^	0030	/08	1333	
		ĺ]	ļ	WQYRQLSALHRAPRPTRPDKARRLGYKAKQ
1			1			GY/VYIYIGFVFAVIYRIRVRRGGRKRPVPKG
		!				ATYGKPVHHGVNQLKFARSLQSVAEERAGR
]	HCGALRVLNSYWVGEDSTYKFFEVILIDPFHK
		ļ	1		1	AIRRNPDTQWITKPVHKHREMRGLTSAGRKS
						RGLGKGHKFHHTIGGSRRAAWRRRNTLQLH
						RYR
814	2164	A	6635	201	1705	KGTEMNKSRWQSRRRHGRRSHQQNPWFRLR
						DSEDRSDSRAAQPAHDSGHGDDESPSTSSGT
				ļ	1	AGTSSVPELPGFYFDPEKKRYFRLLPGHNNCN
		ł	1	1	ł	PLTKESIRQKEMESKRLRLLQEEDRRKKIARM
1		l		[1	GFNASSMLRKSQLGFLNVTNYCHLAHELRLS
		1		I		CMERKKVQIRSMDPSALASDRFNLILADTNS
		Ì	[ĺ	1	DRLFTVNDVTVGGSKYGIINLQSLKTPTLKVF
		1		1		MHENLYFTNRKV\NSVCWASLNHLDSHILLC
		1		1	1	LMGLAETPGCATLLPASLFVNSHPAGIDRPG\
					1	MLCSFRIPGAWSCAWSLNIQANNCFSTGLSR
		ł	1	l	1	RVLLTNVVTGHRQSFGTNSDVLAQQFALMA
]		}	PLLFNGCRSGEIFAIDLRCGNQGKGWKATRLF
. 		Ì	1	l	1	HDSAVTSVRILQDEQYLMASDMAGKIKLWD
		1		1	f	LRTTKCVRQYEGHVNEYAYLPLHVHEEEGIL
		1		1		VAVGQDCYTRIWSLHDARLLRTIPSPYPASKA
		l		1	1	DIPSVAFSSRLGGSRGAPGLLMAVGQDLYCY
		l				SYS
815	2165	A	6643	659	3282	NKNILEVPSARTTRIMGDHLDLLLGVVLMAG
	-100	l . .	5575	1	3232	PVFGIPSCSFDGRIAFYRFCNLTQVPQVLNTTE
]		1	1	RLLLSFNYIRTVTASSFPFLEQLQLLELGSQYT
[l		ļ		
		l	ļ	1	1	PLTIDKEAFRNLPNLRILDLGSSKIYFLHPDAF
		l				QGLFHLFELRLYFCGLSDAVLKDGYFRNLKA
		1	}		1	LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS
				,	ř .	NQIFLVCEHELEPLQGKTLSFFSLAANSLYSR
ľ		l				
						VSVDWGKCMNPFRNMVLEILDVSGNGWTV
				<u> </u>		

NO. of NO. of NO. of NO. of NO. of NO. of No	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nuchec de cotide sequence unoc							
Sequence USSN Sequence 1949/96 Corresponding to first amino acid residue of peptide sequence Pep		_	1200	1	,	ſ	F=Phenylalanine, G=Glycine, H=Histidine,
uence Mence Go/495 Corresponding ng to first and said residue of peptide sequence peptide sequence							
uence 914 mg to first amono acid residue of peptide peptide sequence TThreenine, VValine, VVinco, VVinc					1		
amino acid residue of sequence peptide sequence sequence sequence sequence sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence sequen	-			1			
residue of peptide sequence Papsible nucleotide deliciton, "possible nucleotide deliciton, "possible nucleotide deliciton, "possible nucleotide dissertion RYPETILKDLKVINLAVNKINKIADEAFYGLD	June		ļ	1	. –		
### Possible nucleotide deletion, **possible nucleotide insertion nucleotide insertion in nucleotide insertion in the property of the property							
micleotide insertion RVEPILLOSIEVINLAYNKINKIADEAFYGLD NILQVILNLSYNLLGELYSSNFYGLEKVAYIDL NILQVILNLSYNLLGELYSSNFYGLEKVAYIDL NILQVILNLSYNLLGELYSSNFYGLEKVAYIDL NILDVILLIAVINLAYNLIABLIDISNIALTH FIPSIPDIELSONKLVTLEKINLTANLHLISENS LENLOLLYLLINVINSIPPGVESHLTALR GI.SINSINRLTVLSINDLPANLEILDISNINQLL APPIPOVYSLSVLOITINIFICECELSTIFINVI. APPIPOVYSLSVLOITINIFICECELSTIFINVI. NHTNYTIAGPPADIYCYYPDSLSOVSLESISTS GCDEEVILSILKSIS, FIVCTVTLITLITLITUT TERRGFCFICYKTAQRLVKKOHPQGTEPDMY KYDAYLCFSSKDFTVVONALLKHIDTOYSD QNRFNLCFERRDFVRGENRRANIQDAIWNSR KYDLVLVSRHERLDBWCLAFSYNAGCELSDI. NSALIMVVVGSLSQYQLMKHOSIRGFVQKQQ YLRWPEDLQDVGWFHIALSQQLKKEKKK KDNNIPLQTVATIS BIAGAGRAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1				1	ooq	
RIVETILDISVILIGATE SSINYGLENYATIDL NQVINLSYLLIGATE SSINYGLENYATIDL QKNHHAIIQDOTREFLEKIQITLDIRDNALITH FPSIPIPIES GONKUTYLENDI.TALIHI, SENR LENLDILYFLIRVPH.QILLIAQNIRFSSCSGOD TSENPSLEQUE GERMAQLA WETELCWDVF EGL SHLQVI.YINSINYI.NSLPPGVFSHI.TALR GI.SINSNRLTVI.SINDLPANLELIDISRNQLL APPIDVFVSI.SVLDITINNEFICECELSTFINVI. NHTINVTIAGPPADIVCYPDIS.GSVSLSI.SIST GCDEEVILKSI.KSS.IFVCTVTI.TLFLMTILLTY TERREFCETICYTA.TOR.VFKDHPGOTEPDMY KYDAYL.CESREDYFOSDRANANIQOAUMNER KIVCLVSRHFLBDGWCLEASYAQCRCLSDL NSLLMAVWOSI.SQVJLMKHOSIRGFVOKOQ YLRWPEDLQDVGWFLHKLSQQILKKEKK KIVCLVSRHFLBDGWCLEASYAQCRCLSDL NSLLMAVWOSI.SQVJLMKHOSIRGFVOKOQ YLRWPEDLQDVGWFLHKLSQQILKKEKK KONNIPL.QTVATIS 816 2166 A 6646 I 3811 RDRAGVRPAGKQHA.AAAFTDVGGDEPWDS GNTQLPYRNYKANAMFGAGDEDDTDFLSPS GOARLASLFCLDQAAAGHONEFTOYTARKQP KKQQCTAATONQATRKTAPATMSTPTLLVAT AVBAYBYTNGQYYVKQKGAAVLGHTTR EYRILLYISQQQPVTVARIHVNELMVRPINNY STFYDDGQNWSMRHESEKAAVEFNKQVCIA KCNSTSILDOPQQPVTVARIHVNELMVRPINNY STFYDDGQNWSMRHESEKAAVEFNKQVCIA KCNSTSILDOPQROVENSTANKDKLILIK LGGGEVKKQWEDGMLGMKKGGKRLLIVPPA CAVGSGGMISVAGCHQAVGVPSTANKDKLILIK LGGGEVKKQWEDGMLGMKKGGKRLLIVPPA CAVGSGGMISVAGCHQAVGVPSTANKDKLILIK LGGGEVKKQWEDGMLGMKKGGKRLLIVPPA CAVGSGGMISVAGCHQAVTSQVLQLANTSD AVKAKLISRAKMQQPMB-PIPPQLDSNDSBL EDVNTLQGGGQPVYTSQVPSLQANTSQV VSPFTSIPFKSGGPJALRTKSNLSEQLANTSD AVKAKLISRAKMQQPMB-PIPPQLDSNDSBL EDVNTLQGGGQPVYTSQVPSLQANTSQV VSPTSIPFKSGGPJALRTKSNLSEQLANTSD AVKAKLISRAKMQVAAPPQASAVTSQU PVRPLYPAPLSQPHPQGGGOMASFLMTEAR QHNTEBKMASVXANGMISKELERNNENGRYYEGS NLMMEKKNNSQQTATENTQAKVARAQQP SQCAQGRAYQQKUCKAYAPQAAAVTSQU PVRPLYPAPLSQPHPQGGGOMASFLMTEAR QHNTEBKMASVXANGMISKELERNNENGRYYEGS NLMMEKKNNSQQTATENTQAKVARAQQP SQCAQGRAYQQKUCKAYAPQAAACQLS LVQAELQTQWBAKCEHLLASAKDEHQQVQ CECCAQGAYQQKULKKTNSTIOQAAAGQLS LVQAELQTQWBAKCEHLLASAKDEHQQVQ CECCAQGAYQQKULKKTNSTIOQAHRREP ESPMYPSEQVVERAVPLPQALTTSQDGHRR KGDSEAALSEKUGSLEPHCSCPIRAVLGAP TSIPPPLGPVSMDSECESLAASPMAAKPDD PSGKVCVCRAVAPQDQCSKEESSSEEGE EKAERFRRPSQGGSASSGOQPALNERP ESPMYPSEQVVERAVPLPQALTTSDDGHRR KGDSEAALSEKUGSLEPHCSCPIRAVLGAP TSIPPPLGPVSMDSECESLAASPMAAKCPDD PSGKVCVCRAVAPQDQCSKEESSSEEGE EKAERFR	j						
NILOVILNISYNILLGELYSSINYGLFKVAYIDL (NCHMAIDODTREELKLOTILLINDNALTTH FIRSIPDIFLSONKLVTLEKINLTANLHLISENS LENDLILYFLLRVPH, QUILLINONISPSCSGOD TPSENPSLEQLF.GENMLQLAWETELCWDVY GLISHLQVLYILMIVNISPPGVENSHLTALR GLSINSPRLTVLSHNDLPANLEILDISRNOLL APPEDVFVSISVAUTHINFEICCELSTFINVL APPEDVFVSISVAUTHINFEICCELSTFINVL APPEDVFVSISVAUTHINFEICCELSTFINVL APPEDVFVSISVAUTHINFEICCELSTFINVL APPEDVFVSISVAUTHINFEICCELSTFINVL APPEDVFVSISVAUTHINFEICCELSTFINVL APPEDVFVSISVAUTHINFEICCELSTFINVL APPEDVFVSISVAUTHINFEICCELSTFINVL APPEDVFVSISVAUTHINFEICCELSTFINVL APPEDVFVSISVAUTHINFEICCELSTFINVL APPEDVFVSISVAUTHINFEICLANT KVCLVSRHFIRDGWCLEARSYAQGRCLSDL NSALIMVVOSISQOJAMKORGROFVOKO YTRWFEDLQDWGWFLHKISQOJLKKEKSKK KONNPLOTPATIS SIG 2166 A 6646 I 3811 RDRAGGVRFAGQHAAAAFYDVGGDEPWDS GNOLPFRNVKANAMFOAGDEDDTDLSPS GGARLASLFGLDQAAAGHONEFFOYTARKOP KKQQGTAATGNQATATAPATMSTPILLVAT AVHAYRYTNGQYVVCQKFGAAVLGNHTTIR EYRILLYISQQOPVTAARHONEFHOYTARKOP KKQQGTAATGNQATATAPATMSTPILLVAT AVHAYRYTNGGYVKQGKFGAAVLGNHTTIR EYRILLYISQQOPVTAARHONEFHOYTARKOP KKQGGTAATGNQATATAPATMSTPILLVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTIR EYRILLYISQQOPVTAARHONEVORSLE VAYTGWLFONHVLQQVFDSTANKOKLLIRLK LGSGKVKGWBDGMGAMKGGGRALLIVPA CAYGSEGVIGWTQATDSILVFEVEVRRVKA KDSGSGGHISVSSRDAAPSPIPGADNLSADPV VSPTTSIPFKSGGFALRTKSNSLSQLANTSPIP AVKAKLISRMAKMOQPM-PLIPPQLDSNDSEL EDVITLQGGGOPVTTSVOPSLQFAHPALFO MTSQAPPSVTGLQAFSAAMOVSLUSHBAA VSGNAQSFQYYAGMQAYAYPQASAVTSOLQ PVFPLYPAPLSQPPHFGGGGDMASFLMTARA QHNTERMAVSKVADRMDHLMTKVEELQKK SAGNSALIPBMSVTHETSMMSMIQQENER LKQELLKSNRIEGNDKISLIERNGRYPGOS NLAMERKRNISLGTARETSMMSMIQQENER LKQELLKSNRIEGNDKISLIERNGRYPGOS NLAMERKRNISLGTARETSMMSMIQQENER LKQELLKSNRIEGNDKISLIERNGRYPGOS NLAMERKRNISLGTARETSMMSMIQQENER LKQELLKSNRIEGNDKISLIERNGRYPGOS NLAMERKRNISLGTARETSMMSMIQQENER LKQELLKSNRIEGNDKISLERNGRYGATALAGAPCH QTOLTSKLETDLAGGLTXVOAKLSELGET SEQAQSKYKSKONFQLIKATSLGEELDI RKSTYQEELDILAKRYRYSTOQAACALALGA DFISK-KYKINNYQPGSRFARELEESSEEDEE EKAERFRRFSQGSASSGOQPALNERP ESPMYSEQVVERAVPLPQAALTTSODGHRR KGDSBAALSEKUGSSPELSCESSSEEDEE EKAERFRRFSQGSASSSGOPQAALNERP ESPMYSEQVVERAVPLPQAALTSODGHRR KGDSBAALSEKUGSSEDPE			-		2042000		<u> </u>
OKNHHAJIODOTTKFLEKUTILDLADNALTHIH FPSIPDIFLSORIKUTIKRNITANILHISEN LENLDILYFLLRVPH,QILLIAONRESSCSGOD TPSENPSLEGUT GERMALQIA WETELCWDVF EGL SHLQVLYJNBINYLNSLPPGYFSHLTALR GLSLNSNLTVLSHDDLPANLELDISRNOLL APNPDVFVSLSVLDITHREFICECELSTFINWL NHTNYTAGPADIYCVYPDSLSGVSLSSLSTE GCDESEVLSLKPSLTFVCTVTLTILFMILTVY TKFRGFCFICYKTAQRLVFKUDTGYSSDVSLSSLSTE GCDESEVLSLKPSLFFVCTVTLTLTHATILTV TKFRGFCFICYKTAQRLVFKUDTGYSSDVSLSSLSTE GCDESEVLSLKPSLFFVCTVTLTLTHATILTV TKFRGFCFICYKTAQRLVFKUDTGYSSDVSLSSLSTE GCDESEVLSLKPSLFFVCTVTLTHATILTV TKFRGFCFICYKTAQRLVFKUDTGYSSDVSLSSLSTE AVYDVLCSSKDFTVVQVALLKHIDTGYSD QNREPILCPEERDFYPGENRYANIQDAJWNSR KYDAVLCSSKDFTVVQVALLKHIDTGYSD QNREPILCPEERDFYPGENRYANIQDAJWNSR KYDAVLCSSKDFTVVQVALLKHIDTGYSD QNREPILCPEERDFYPGENRYANIQDAJWNSR KYDCLYSHIPLDQVGULARSYNQCALSSIS NSALIMVVVGSLSQVQLMKHGSIRGFVQKOQ YLRWPEDLQDVGWHKHGSQGLKKEKKK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGRYBAGKGHAAAAFYDVGGDRPWDS GRARLASIFGLDQAAAGHONEFFOYTAFKQP KKQQGTAATGNQATFKTAPATMSTPTILVAT AVYAYAYYTNGQVVGGKGAAALGHTTR EYRILLYISQQQPYTVARHAVFELMVRPNNY STFYDDDQRQNNSWIRFSESKAAVEFNLQVCIA KCNSTSSLDAVLSQDLVAADGAVCRGHCAALLGHTTR EYRILLYISQQQPYTVARHAVFELMVRPNNY STFYDDDQRQNNSWIRFSESKAAVEFNLQVCIA KCNSTSSLDAVLSQDLVAADGAVCRGFCALLIAVRPNY VSPFTSIFFKSGFALRTKSNSLSEQLAINTSPTLLVAT CAVVSEGOVIGWTQATDSLVFFEVERRVKLA KDSGSDGISVSSRDSAAPSPFPGADILSADFY VSPFTSIFFKSGFALRTKSNSLSEQLAINTSDR AVXAKLISRBAKMQQMAJYAYQASAVTSQLQ MTSQAPQFSVTGLQAFSAALMQVSLDSHSA VSGNAQSFQFYAGMQAYAYPQASAVTSQLQ PVPRLYAPALSQPFHQGSGDMASFLMTEAR QHNTERMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTURFTSMMSNIQRIQENER LKQELEKSNTIEGRONIVGLEKKTSLEELTDI RVKKSSLEKNLSERKKSAQERSQCSAASASQQAPLNRERP ESPMYSEQVVERAVPLPQAALTTSQDGRRR LKQELEKSNTIEGRONIVGLEKKTSUSEELTDI RVKKSSLEKNLSERKKSAQERSQCVEAAAS DPSEKWKLMNQVPQSLRREFTLEESYNGRTI LGTMNTIKMVTLQLLAQQFQCKEESSSEEE ERAERFRRTSQCSAASSGOQPAFLNRERP ESPMYPSEQVVERAVPLPQAALTTSQDGRRR KGDSBAALSSKLKOSLEPHLSCTSRRVLGFP TSIPPEPLGPVSMDSCCESLAASPMAAKDPDN PSGKWCVERVAPOPQLQESTETISLTSDPFE GDPLALGFESFGPPQPQLKKDDVTSSTOPHK ELSSTFAGASTVAGAALRSHH9GRSSLSGDEE GDFLALGFESFOPPQPQLKSTRUSTSDPFE GDPLALGFESFOPPQPQLKKDDVTSSTOPHK ELSS	{ :		ĺ				
FIPSIPDIFI, SGNKL, VTLPKINLTANLIHLISEN LENLDIL, YFILA VPHI-OLLIN (MORNESSCSIG) TPSENPSI.EQLF.I.GENM.QLAWETEL.CWP) EGI.SHI, QVI, YIMNYINSI, POYVSHITALR GLSIANSNRI, TVLSHNDLANLEILDISRNQLL APPPDYFVSISLYDITHKREICESTRIWIL APPPDYFVSISLYDITHKREICESTRIWIL NHTHVTIAOPPADIYCVYPDSI.SGVSLESTRIWIL APPPDYFVSISLYDITHKREICESTRIWIL NHTHVTIAOPPADIYCVYPDSI.SGVSLESTRIWIL KYDAYALCPSSKDFTUPQNALLIKIDTOYSD QNRFNL.CFEERDFYPGENRPANIQDAJWNSR KYCLVSRHFLRGWCLEAFSY, AGRCLSD. NSALMAVVVGSI.SQYQLMRHQSIRGFVQKOQ YLRWPEDLQDVGWFLHKI.SQQILKKEKEKK KDNNIPLCTYATIS 816 2166 A 6646 1 3811 RDRAGVPFAGKQHAAAAFDVGGDAFWDS GNTQLPFRNPYKANAMFGAGDEDDTDFILSPS GGARLASLFGLDQAAAGHGNEFFQYTAFW KKQGTAATGNQATKTAPATHTILVAT AVHAYRYTNOQYYKQAKGAAAVLONHTIX EYRILLYISQQQPVTVABHINVFERWPNNY STFYDDQRQNWSIMFSSEKAAVENKQCUL KCNSTSSLDAVI.SQDLIVADQPAVGVOGDLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEGDMLGMKKGGAAVLONHTIX CAVGSEGVIGWTQATDSILVFEVEVERVKUL LGSGKVIKGWEGDMLGMKKGGAAVLONHTIX CAVGSEGVIGWTQATDSILVFEVEVERVKUL KLKLLINGSQCGCONTONTONTONTONTONTONTONTONTONTONTONTONTO	ļ		i				
LENLDILYFLRYPH.OLILINQNRFSSCSORD TYSENPSI.GOLF, GERMMQLAWETELCWDVF EGI.SH.QVI.YIMHYINSI.PPGVFSH.TAL GL.SINSNRI.TVI.SHNDI.PANLEDISRNQIL APNPDVFVSI.SVI.DITHRKFICECELSTFINQL APNPDVFVSI.SVI.DITHRKFICECELSTFINQL HITNYTIAOPPADIVCYYPDOLSGYSI.FI.STE GCDEEVI.KSI.KFSI.FIVCTVTLTI.FI.MTILTV TKRRGFCFICYKTAQRI.VFKDHPQGTEPPMV KYAPAYLCPSSKDFTWQNALI.KILDITQYSD QNRPNI.CFERDFVGENRPANIQDAIWNS KIVCLVSRHFI.RGWCLEAFSY AQGRCLSDL NSALIMVVVGSI.SQYQLMKHQSIRGFVQKO YI.KWPEDLQDVGWFHIKI.SQQII.KKEKKK KDNNIPLQTIVATIS BI6 2166 A 6646 I 3811 RDRRGVRPAGKQHAAAAFYDVGGDRPWDS GNTQLPFRNPVKANAMGAODEDDTDIFLS; GGARLASI.FGI.DQAAAOHGHFTQYTAFKQ KKGQGTAATONQATPKTAPATMSTFITLVA AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPVTVARHIVNFELMVRPNIV STFYDDQRQNWSINFSEKAAVLGNHTTR EYRILLYISQQQPVTVARHIVNFELMVRPNIV STFYDDQRQNWSINFSEKAAVLGNHTTR EYRILLYISQQQPVTVARHIVNFELMVRPNIV STFYDDQRQNWSINFSEKAAVLGNHTTR EYRILLYISQQQPVTVARHIVNFELMVRPNIV STFYDDQRQNWSINFSEKAAVLGNHTX AVHAYRYTNGQYTAATOSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTGATSCALACT CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATDSLUTYERIV CAVGSEQVIGWTQATATOSLOTARI CAVGSEQVIGWTQATLARQCKTKVQAKLSELQET SEQAQSKFKSEKQNRKQLEKVTSLEELELTI QMQLTSLKETIDLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLEKVTSLEELELTI QMQLTSLKETIDLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLEKVTSLEELETIL QMQLTSLKETIDLRGQLTKVQAKLSELQET CACQRDAYQQKLVQLQEKSEVSTACLALAQ QTTALTKQMCGURKELELENGRYVEG SEMPVESEQVYEAVPLIPQALTTSQCGGREDDEI RKSYQEELDKRRQLIKKTRYSTDOAAABOLS LVQAGLTOTOWARCEHLELDANGSQCKAAAS DPSEKVKKIMOQYPOSLRREFELEESYNGELGE EKAEERPRRFSQCOSASASSQQALALAS ADPSEKVKKIMOQYPOSLRREFELEESYNGE ESPMVPSEQVYEAVPLIPQALTTSQCGGREDDEI FKSTQCAVGARATSSHINGRSLSGDEE EKAEERPRRFSQCOSASASSQQALALSCHINGDS SCORD STREED STRUTTSTOTT CATCHATACT CATCHATACT CATCHATACT CATCHAT		1					, , , , , , , , , , , , , , , , , , ,
TPSENPSLEQLFLGENMULAWETELCWDVF EGLSHLQVLYLINHVINSIPPOSHLTALR GLSLMSNRITVLSHNDLPANLELDISRNOLL APPPDVFVSLSVLDTINKFE(CECLSTFINW). HTMVTTAGPPADIYCVYDDLSGVSLFSLSTE GCDEELVLSILKFSLFIVCTVTLTLFLIMTLLTV TEKRGPCEICYKTAGRUFVENDPGENPANHOLDAWSS NALIMVVVGSLSVQLMKHGPTOREPDMY KYDAYLCFSSEDFTWQNALLERLDITQVSJ QNERPINLCFEERDFYPGENPANHOLDAWSS KIVCLVSHFLRDGWCLAFSYAGGRCLSHD NSALIMVVVGSLSYQLMKHGSIGFVQKOQ YLRWPEDLQDVGWFLHKLSQQILKKEKEK KDNIPPLQTVATIS 816 2166 A 6646 I 3811 RDRAGVPFAGKQHAAAFYDVGGDRPWD KKGQGTAATGNAATFYDVGGDRPWD KKGQGTAATGNAATFYDVGGDRPWD KKGQGTAATGNAATFYDVGGDRPWD KKGQGTAATGNAATFYDTHTLVAT AVHAYRYTNOQYVKQGKFGAAAFLCNHTTR EYRLLYISQQPVTVARHIVPTAMTSTPILVAT AVHAYRYTNOQYVKQGKFGAAAVLCNHTTR EYRLLYISQQQPVTVARHIVPTAMTSTPILVAT KCNSTSSLDAVLSQDLIVABGPAVEVOBSL VAYTGWFGNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVABGPAVEVOBSL VAYTGWFGNWTQATDSSLVFVERVIKA KDSGSDGHSVSSRDSAAPSPPGADNLSADPV VSPPTSIPFKSGEFALRTISNSLSQLANTSPD AVKAKLISRMAKMGQPMJPLPPQLDSDSDE EDVNTLQGGGGPVTSVOPSLQAHFALPQ MTSQAPQFSVTGLQAFSAALMQVSSLDSHA VSGNAQSFQPVAGMQAYAPPOSAVTSQLA VSRNLSSRAAMGSPPHGASALTMVSLLGANTSPD AVKAKLISRMAKMGQPMJPLPPQLDSDSDE EDVNTLQGGGGPVTTSVOPSLQAHFALPQ MTSQAPQFSVTGLQAFSAALMQVSSLDSHA VSGNAQSFQPVAGMQAYAPPOSAVTSQLA QHTERMAVSK VADKMPHLMVELQQKB SAGNSMLIPSMSVTMETSMIMSNIQRIQDER LKQEILEKSNRIEGNDKISELIERNQRYVEGE NAMMERNNSLQTATENTQARVHAEQEKA KVTELLAAATAQVSHLQLKNTAHQKKSTLDSH SGQAQSKFKSEKQNRKQLELKVTSLEEELIT QMQLTSLKETIDLRGQLIKVGAKLSELGET SGQAQSKFKSEKQNRKQLELKVTSLEEELIT QMQLTSLKETIDLRGQLIKVGAKLSELGET SGQAQSKFKSEKQNRKQLELKVTSLEEELIT QMQLTSLKETIDLRGQLIKVGAKLSELGET SGQAQSKFKSEKQNRKQLELKVTSLEEDLDL RVSKYGEELDKLRQLIKKRYSTDQAAASQ QTTALTKQNEGQKEELSSSSEEEE LKAEERRRPSGCSASASGQPQADLNER ESMYPSEQVVEALVPPQALTTSQDGHRR KGDSEABALSEIKGGSLPPELSGYFRRVLGPP TSIPPEBLGRVSMGSCEESLSASSGOPQALANSP PSEKVKKIMNQVPQSLRREFELESYNGT LGTMNTIKMVTQLLANQCGEGEESSSEEEE EKAEERRRPSGCGASASSGOPQADLREF ESMAVPSEGVVEALAYBLAGRESLSLTSDPE GDPLALGPESPGEPQPQLKKDDVTSSTGPD PSGKVKKIMNQVPQSLRREFELESYNGT LGTMNTIKMVTQLALREFELESTNGTSDPE GDPLALGPESPGEPQPQLKDDVTSSTGPD PSGKVKKIMNQVPQSLRREFELESTNGTSDPE GDPLALGPESPGEPQPQLKKDDVTSSTGPD BCLFKGA	1	ŀ					
EGLSHL, QVL, YLMHNYLINŠ, PPGVPSHI ALIG GLSHL, QVL, YLMHNYLINŠ, PPGVPSHSHOLAL GLSHNSNI, TVLSHNDL ANALELDISRNQLL APPPDVPSUS VLDHTHKFICCECL STFIND AVAYLCPS KDR, PWC, PVCTUT, I.F. IMITLIV TKFRGPCEICYKTAQRL VFKDHPQGTEPPM KYDAYLCPS KDFTWQNALLKHLDTQYSD QNRPNLCFEERDFYPQENRPANIQDAIWNS KIVCL VS SHHI RDGWCLEAFSY AQGRCLSDL NSALIMVVVGSLSQYQLMKHQSIRGPVQKG YLKWPEDLQDVGWFLHKLSQQILKKEKKK KDNNIPLQTVATIS BERAGVRPAGKQHAAAAFDWGGDFPWDS GNTOLPPRAPV&ANAMFGAGDEDDTDFLSPS GGARLASLFGLQAAAAGHGNEFFQVYAFKQ KKQQTAATONQATPKTAPATMSTPTILVAT AVHAYRYTNQQYVKQGKFGAVLGNHTTIL EYKILLYISQQQVTVARHIVNELMVRPNNY STFYDDQRQNWSIMFESEKAAVENKQVCIA KCNSTSSLDAVLSQDLIVADQPAVCOSLE VAYTGWLFQNHVLGQVFDSTANKDKLLIVRA LGSGKVIKGWEDGMGMKKKGCKLLIVPPA CAYGSEGVIGWTQATDSILVFEVEVRVRVLA LGSGKVIKGWEDGMGMKKKGCKLLIVPPA CAYGSEGVIGWTQATDSILVFEVEVRVRVLA KCNSTSSLDAVLSQDLIVADQPAGDNISADPV VSPPTSIPPKSGEPALRTXSNSLSEQLANTSPD AVKAKLISRMAKMGQMDLIPDGSNDSEI EDVNTLQGGQQVVTPSVQPSLQRAHPALPQ MTSQAPQPSVTGLQAFSAALMQVSSLDSHA VSGNAQSFQPT AGMGAYAYPQASAVTSQLQ PVRPLYPALSQPPHFQGSGDMASFLMTEAR QHNTERMAVSKVADKMDHLMYELLQKH SAGNSMLIPSMSVTMETSMIMSNIQRIJGERS LQGGLEKSNIEGEONDKISELERNQRYVEGS NLMMEKRNISLQTATENTQARVLHAEGKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEQNRKOLERNTSLEELITDL RVEKESLEKNLSERKKSAQRESQAEEEIDEI RKSYQEELDKLRQLIKKTRVSTDQAAAGU QTTALTKQNEGQHKLELENNSKNGSVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGTL LGTMMTIKMYTLQLLQCQEGKEESSSEEEE EKABERPRPSQEGSASASSQPQALNER LGGBLEKSNIEGKLELENNSKNGSVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGTL LGTMMTIKMYTLQLLQCQCGKEESSSSEEEE EKABERPRPSQEGSASASSQPQALNER ESPMYSEQVVEAPLPPQALTTSQDGFRR KGDSEAALSERLOGSLPPELSCPSTRALLGTTSDPEE GDPLALGPESPGEPQPQLKKDDVTSSTGPHP ESPMYSEQVVEAPLPPQALTTSQDGFRR KGDSEAALSERLOGSLPPELSCPSTRALGTBDPE ESPMYSEQVVEAPLPPQALTTSQDGFRR EGDPLALGPESPGEPQEPQLKKDDVTSSTGPHP ELSTEAGSTVAGAALRSPENSGSSLSGDEE EKABERPRPSQEGSASASSQPQALANER DPSEKVKRIMQVFQSLRALGRINGSSLSGDEE EKABERPRPSQEGSAASSQOPQALNER ESPMYSEQVVEAPLPPQALTTSQDGFRR KGDSEAASSGCPQALTSCREET ELSTEAGSTVAGAALRSPENSGSSLSGDEE EKABERPRPSQEGSEEDEEDEVSMKGR PPTTELGDDDDDDDWLG	ł	}	1	}	}		* *
GLSLNSNRLTVLSHNDLPANLELIDISRNOLL APPNDVYSLSVLDITHRNRFICECELSTSTINML NHTHNYTHAGPPADIYCYPPDSLSGVSLSSTS GCDEEDYLSKLSKSLFVCYTALTI.FLMTLTV TKFRGFCFCVXTAQRLVFKDHPQGTEPDMY KYDAYLCFSKDFTWVQNALLKHLDTQYSD QNRPNLCFERRDFVPGENFRANIQDAIWNSR KIVCLVSRHFLRDGWCLEAFSYAQGRCLSDL NSALIMVVVGSLSQYGLMHGSRGFVQKQQ YLRWPEDLQDVGWTHHGLSQQLKKEKK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGVRFAGKQHAAAAFYDVGGDRFWDS GNALIMSVVGCKGKGAAAAAFYDVGGDRFWDS GNALIMSTVANTSTELVATIN GOYLQFRIPYKANAMFGAGDEDDTDFLSFS GGARLASLFGLDQAAAGHGNEFFCYTAFKQP KKGQOTAATGNQATPKTAFTGYTTAFKQP KKGQOTAATGNQATPKTAFTGYTAFKQP KKGQOTAATGNQATPKTAFTGYTAFKQP KKGQOTAATGNQATPKTAFTGYTAFKQP KKGQOTAATGNQATPKTAFTGYTAFKQP KKGQOTAATGNQATPKTAFTGYTAFKQP KKGQOTAATGNQATPKTAFTGYTAFKQP KKGQOTAATGNQATPKTAFTGYTAFKQP KKGQOTAATGNQATPKTAFTGYTAFKQP CAYGSEGYIGWTQATTSBLLYEVEVERVKKIA KDSGSDGHSVSSRDSAAFSPIPGADNLSADFV VSPPTSIPFKSGPFALTKKSNLSEQLAINTSPD AVKAKLISRNAKMGQPMLPLPPQLDSNDSEI EDVNTLQGGGQPVVTTSVQPSLQFAHFALPQ MTSQAPQFSVTGLQAFSAALMQVSSLDSISA VSGRAQSFGPYAGMQATAYPQASAVTSQLQ PYRPLYPAPLSQPPHFGGSGDMASFLMTEAR QHNTEIRMAVSLVADKMDHTKYVELGKF SAGNSMLPSMSVTARETSMIMSNIQRIIGENER LKQEILEKSNREEQNBKSELERNGRYTEGS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAFQKKETPL QMQLTESLKENTLESRRKSVAGCRSQAEEGIDEI RKSYQEELDKLRQLLKKTTKYSTDQAAAGLS LVQAELGTQWFAKCEHLLASAADEHLQQVQ EVCAQRDAYQQKLVQLGEKSSVCERE LGTMTNTTLANTAGLERNAKLSERLKKSAGCRSQAEEGIDEI RKSYQEELDKLRGLLKKTKTSTDGAAAGLS LVQAELGTQWFAKCEHLLASAADEHLQQV EVCAQRDAYQQKLVQLGEKSSVCERE LGTMTNTTLANTAGLERNAKLPGP SEPRAPPSGGSAASSGQPQAPLNERP ESPMYPSEGVYEAVALPOPGLESSTRISLSTDPEE EKAEERRRRSGEGSAASSGROQAPLNERP ESPMYPSEGVYEARVAPLOPGLESSTRISLTSDPEE GDPLALGPESFGEPOPPQLKKDDVTSSTOPHK ELSSTEAGSTYAGAALARVSNSCSSGDEE DELFKGATLKALAPKAQPEEDEDEVSMKGR PPPTTPLFGDDDDDDDUBLG				}			,
APNPDVFVSLSVLDITINTRICECELSTITIWU NHTNVTIAGPADITYCVYPGSSVSJESISSS GCDEEEVLKSLKFSLFVCTVTLIT.HT.MILTV TKFRGFCFCVXTAQRLVFKDHPQGTEPDMY KYDAYLCFSSKDFTWVQNALLKHLDTQYSD QNRFNLCFEERDFVPGENRPANIQDAIWNSR KIVCLVSRIFLERDGWCLEAFYAQGRCLSDL NSALIMVVVGSLSQVQLMKHQSIRGFVQKQQ YLRWFDLQDVGWFLHALSQQLKKEKEKK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGVRPAGKQHAAAAFYDVGGDRFWDS GNTQLPPRNPVKANAAMFGAGDEDDTDLSPS GGARLASLFGLDQAAAGHGFQVTAPKOP KKGQGTAATGNQATFKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPVTVARHVNFELMVRFNNY STFYDDQRQNWSIMFSSKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGAVEVGOSLE VAYTOWLPQNIPULQQVFDSTANKDKLRLK LGSGKVIKG WEDOMLGMKKGGKRLLIVPPA CAVGSEGGIGWTQATDSILVFEVEXFXVIA KDSGSDGHSVSSSBAAAFSPTGADNLSADFV VSPPTSBFFKSGEPALRTKSNSLSEQLANTSPD AVKARLISRAMKGGPAPPOLDSIDNSEI EDVNTLQGGGOPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHA VSGNAQSTQPYAGMQAYAPQASAVTSGLQ PVRRLYPAPLSQPHPGGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTAMETSMIMSNIGNERE LKQEILEKSNRIEEONDKISELIERNQRYVEGS NLMMERKNNSLQTATENTOARVLHAGGEA KVTEELAAATAQVSHLQLKMTAPGKSET LKQEILEKSNRIEEONDKISELIERNQRYVEGS NLMMERKNNSLQTATENTOARVLHAGGEA KVTEELAAATAQVSHLQLKMTAPGKSET LKQEILEKSNRIEEONDKISELIERNQRYVEGS NLMMERKNNSLQTATENTOARVLHAGGEA KVTEELAAATAQVSHLQLKMTAPGKSET LKQELEKSNRIEEONDKISELIERNQRYVEGS NLMMERKNNSLQTATENTOARVLHAGGEA KVTEELAAATAQVSHLQLKMTAPGKSET LGTMNTIKMVTQLLLORGCESSSCEEEL CKSERRERSPSGDSASASSGQPQAAHREP ESPMYPSGQVEAVALPQALTISTQDGHRR KGDSEAEALSEIKDGSLPSELSCPSHRVLGP TSIPPPPLGPVGMSCEELLAAARAACLS LYQAELGTQWEAKCEHLLAAAACLS LGTMNTIKMVTLQLLORGCESSESSEEEL EKAEERRRFRSGEGSASASSGQPQAHREPP ESPMYPSGGVEAVAPOAATCJDP PSGKVCVREVAPOGPLQSSTRISLTSDQFE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALARSSALGSDFSSLSGDEE EKAEERRRFRSGEGSASASSGQPQAHAAKVPDN PSGKVCVREVAPOEPLQSSTTISSTGPHK ELSSTEAGSTVAGAALCRSTATISTSDQFE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALCRSTATISTSDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALCRSTSTGPHK ELSSTEAGSTVAGAALCRSTSTGPHK ELSSTEAGSTVAGAALCRSTSTGPHK ELSSTEAGSTVAGAALCRSTSTGPHK ELSSTEAGSTVAGAALCRSTA	i		į				1
NHTHVTTAGPPADIYCYPPDSLGGVSLSISLSTSTE GCDEEGVLSLKSS.HVCVTVITLIF.MTLITV TKFRGFCFICYKTAQRLVFKDHPQGTEPDMY KYDAYLCFSKDFTWVQNALLKHLDTQYSD QNRFNLCFERDFYPGEBRFANNQDAIWNSR KIVCLVSRIFILRDGWCLEAFSYAQGRCLSDL NSALMYVVOSLSQVQLMKHQSIRGFVQKQQ YLRWPEDLQDVGWFLHKLSQQILKKEKEK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 GDRAGVRPAGKQHAAAAFYDVGGDRWDS GNTQLPFRNYRVANAMFQADEDDTDFLSPS GGARLASLFGLQAAAGHGNEFFQYTAPKQP KKQQTTAATGNQATFKTAPATMSTPTILVAT AVHAYRYTINGQVVKQGKGAAAAFYDVGGDRWDS GARLASLFGLQQAAAGHGNEFFQYTAPKQP KKQQTTAATGNQATFKTAPATMSTPTILVAT AVHAYRYTINGQVVKQGKGAAAAHUNHTIR EYRILLYISQQQPVTVARHVNFELMWRPNNY STFYDDQRQMVSIMFSESKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDANADKLALIKL LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVEVEVRVKUA. KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPFTSIPFKSGEPALRTKSNSLSQLAINTSD AVKAKLISRMAKMQOPHLPIPPQLDSNDSEI EDVNTLQGGQPVVTTSVQPSLQRAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGRAQSFQPYAGMQATAYPQASAVTQLQ PVRRLYPPALSOPPHGGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVELLQKH SAGNSMLPSMSVTMETSMINSRQUIGQERE LKQELLEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQAVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRQQLTKVQAKLSELQET SEQAQSKFYESKQNRKQLEXTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEDEL RKSYQEELDKLRQLKKTRVSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEDEL RKSYQEELDKLRQLKKTRVSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEDEL RKSYQEELDKLRQLKKTRVSLEEELTDL RVEKESLEKNLSERKKSAQERSQAEEDEL RKSYQELDKLNRQLKTRVSTDGAAAEQLS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEKESSSEEELE EKAEERPRRPSQEQSAASSGQPQAHNERP ESPMYPSGQVEAVPLPPQLATTSODGHRR KGDSBAEALSEIKDGSLPPELSCPSHRICHE EKSTEAGSTVAGAALREFELEESYNGRTI LGTIMNTIKMVTLQLLNQCEKESSSEEELE EKAEERPRRPSQEQSAASSGQPQAHNERP ESPMYPSGQVEAVPLPPQLATTSODGHRR KGDSBAEALSEIKDGSLPPELSCPSHRICHER RKSTYGPFLGKPSMDSCEESSSEEGE EKAEERPRRPSQEQSAASSGQPQAHNERP ESPMYPSGQVFEAVPLPPQLATTSODGHRR KGDSBAEALSEIKDGSLPPELSCPSHRINGRSTOPHK ELSSTEAGSTVAGAALRPSAAKKPDD PSGKVCVREVAPDGPLQESSTELSLTSDPEE GDPLALGPESPGEPPPQLKKDDVTSSTOPHK ELSSTEAGSTVAGAALARPSHGRSSLSGDEE DELFKGATLAALAPKAQPEEDEDEDEVSMKGR PPPTTPLFGDDDDDDDDWLD	1		l				I
GCDEEEVLKSLKFSI-FUVCTVLITLIFLATILITU TKFRGFGFCIVCYTLAQAU VERDIPOGTEPDMY KYDAYLCFSSKDFTWVQNALLKHLDTQYSD QNRPNLCFEERDFVPGENPRANIQDAIWNSR RIVCLVSRIFILRDGWCLEAFYAQGRCLSDL NSALIMVVVGSLSQVQLMKHQSRGFVQKQ YLRWPEDLQDWGWFLHALSQULKKEKEKK KDNNIPLQTVATIS BI6 2166 A 6646 1 3811 RDRAGVRPAGKQHAAAAFYDVGGDRPWDS GNTQLPPRNPVKANAMFGAGDEDDTDLSPS GGARLASLFGLDQAAAGHGNEFCYTAPKQP KKGQGTAATGNQATFKTAPATMSTPITLVAT AVHARYTYNGQVKQGURGHASTPTILVAT AVHARYTYNGQVKQGURGHASTPTILVAT AVHARYTYNGQVKQGURGHASTPTILVAT AVHARYTYNGQVKQUGRGKGKRLLIVPPA STFYDDQRQWWSIMFSSKAAVERNKQVCIA KCNSTSSLDAVLSQDLIVADGAVEVGDSLE VAYTGWLFQNHVLQQVFDSTANKDKLRLK LGSKVKKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGGIGWTQATDSILVFEVVSKYNIA KDSGSDGHSVSSRDAARSPFGADNLSADPV VSPFTSIFFKSGEPALRTKSNSLSQLANTSPD AVKARLISRAMKGGYMLJDPPQLDSNDSEL EDVNTLQGGOGPVYTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAFSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYATPQASAVTSQLQ PVRRLYPAPLSQPHFFGGSGDMASFIMTAR QHNTEIRMAVSKVADKMDHLMTKVELIQKH SAGNSMLESMSVTHETSMMSNIGRIENER LKQEILEKSNRIEEQNDKISELIERNQRYYEGS NLMMEKRNNSLQTATENTQARVLHAGGEA KVTEELAAATQVSHLQLKKTRVSLEELEND RVEKESLEKNLSERKKSAQERSQAEEEIDEI RKSYQEELDKLRIQQLTKVQAKLSELQET SEQAQSKFKSKKOMRYQULGLKKTRVSLDEAAAEQLS LVQAELGTQWBCKKCBHLASAKDEHLQQV EVCAQRDAYQQKLVQLQEKSSVFSACLALQA QITALTKOMPGHIKELEKNKSQMSGVBAAAS DPSEKVKKIMNQVFQSLREFELESPYNGRTI LGTIMMTIKMVTLQLLNQGE(EKSESSEEELE EKAEERTRRFSQEDSASASSGQPQAHAKPDN PSGKWCVREVAPDGPLQSSTRISLTSDGEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTYAGAALCRSTJONDE	į.						
TKFRGFCFICYKTAQRI VFKDHPOGTEPDMY KYDAYLCFSKDFTWVQMALLKHLDTQYSD QNRPNLCFEERDFVPGENRPANIQDAIWNSR KIVCLVSRHFLRDGWCLEAFSYAQGRCSDL NSALMYVVGSLSQYQLMKHQSRGFVQKQQ YLRWPEDLQDVGWHHKLSQQLKKEKK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGVRPAGKQHAAAAFYDVGGDFWDS GNALMSTVAGAQGEDDTDFLSSS GGARLASLFGLDQAAAGHGNEFFQYTAPKQP KKQQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKGAAVLGNHTTR EYRLL YISQQQPVTVARHVNFELMWYRNNY STFYDDQRQNWSIMFSEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVQDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPA CAVGSEGYIGWTQATDSLVEVEVRVKLA KDSGSDGHSVSSRDSAAPSPPGADNLSADPV VSPPTSIPFKSGBAAFLKTKSNSLSQLABNTSPD AVKAKLISRMAKMQQPALPILPPQLDSNDSEL EDVNTLQGGQPVYTSVQPSLQFAHFALPQ MTSQAPQPSYTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRRLYPAPLSQPPHGGSGAMSFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVELQKH SAGNSMLPSMSVTMETSMMSNIQRIQEDERE LKQEILEKSNRIEEQNDKISELIERNQRYVEGS NLMMEKRINNSLQTATENTQARVLHAGEKA KVTEELAAATAQVSHQLKMTAHQKKETEL QMQLTESLKETDLLRQQLTKVQAKLSLQET SEQAQSKFKSKKNAQPRSQAEEDDEL RKSYQEELDKLRQLKKTYSTDQAAAEOLS LVQABLQTQWBAKCEHLLASAKDEHLQQVQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNISQMSGVBAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQCEKSESSEEEEE EKAEERPRRPSQEQSASASSQQVQAHAREPE EKAPERPRRPSQEQSASASSQQVQAHAREPE EKAPERPRRPSQEQSASASSQQVQAHAREPE EKAEERPRRPSQEQSASASSQQQASKTSCESSSEEEE EKAEERPRRPSQEQSASASSQQQABALREPE EKAPERPRRPSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAEERPRRPSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRPSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEQSASASSQQQABALREPE EKAPERPRSQEGSSESSEEEE EKAEERPRRPSQEGSASASSQQQABALREPE EKAPERPRSQUESSTECHESTYTAGDEB BELSTEAGSTVAG	1	1	[[
RYDAYLCFSSEDTFWQNALLKHIDTQYSD	ŀ				ľ		
QNRFNLCTEREDPYPGENRRANIQDAIWNSR KIVCL VSRHFLRDGWCLEAFSYAQGRCLSDL NSALIMVVVGSLSQYQLMKHQSIRGFVQKQQ Y1.RWPEDLQDVGWFLHKLSQQIKKEKEKK KDNNPILQTIVATIS BDRAGVRPAGKQHAAAAFYDVGGDRPWDS GMTQLPPRPYKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGHONEFGYTAFKQP KKGQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYYKQGKFGAAVLGNHTTR EYRILLYISQQQPVTVARIHVNFELMVRPNNY STFYDDQRQNWSIMFESSKAAVEFNRQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGYIGWTQATDSLIVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPTISIPFKSGEPALRITKSNISLSQLAINTSPD AVKAKLISRMAKMGGMPJPLPPQLDSNDSEI EDVNTLQGGQPPVTPSVOPSLQAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHPQGSGDMASFLMTEAR QHNTBERMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIGNIGDENER LKQEILEKSNIEGONDKISELIERNQRYVEGS NLMMEKRNNSLQTATENTQARVHACKETEL QMQLTESLKETDLLRGQACPSQAEEDDEI RKSYQEELDKLRQUKKTYVAGALSLEGLTD RVEKESLEKENISRRKKROGPSQAEEDDEI RKSYQEELDKLRQUKKTRVGALSLEGLTD RVEKESLEKENISRRKKROGPSQAEEDDEI RKSYQEELDKLRQUKKTRVGALSLEGLTD LGTMNTIKMYTLQLLNQGEOEKEESSSEEEE EKAEERPRRPSQEQSASASSQPQAPLNRERP ESPMYPSEQVVELAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCTBRRVLGFP TSIPPEPLGPSPGEEDPGLESSTRLSTGPHK GGDPLALGPSSPGEEPDGLESSTRLSTGPHK BLSSTEAGSTVAGAALRPSHASQRSSLSGDEE DBLIFKGATLKARPKAQREEEDEDEVSMKGR PPPTPLEGDDDDDDDDDUNLG		ļ					
KİVCLUSRHILRIGWCLEAFSYAGRCLSDL NSALIMAVVGSLSQYQLMKHOSIRGPYQKQQ YLRWPEDLQDVGWFLHKLSQQILKKEKEKK KDNNIPLQTVATIS RDRAGVFRAGKQHAAAAFYDVGGDRPWDS GHTQLPPRNPVKANAMFGAGDEDDTDFLSFS GGARLASLFGLDQAAAGHONEFFQYTAPKQP KKGQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTR EYRILLYISQQQPYTVARIHVNFELMVPRNNY STFYDDQRQNWSIMFESEKAAVEFNRQVCIA KCNSTSSLDAVLSQDLVAGGREGAAVLGNHTR EYRILLYISQQQPYTVARIHVNFELMVPRNNY STFYDDQRQNWSIMFESEKAAVEFNRQVCIA KCNSTSSLDAVLSQDLVAGGRVEVGOSLE VAYTGWLFQNHVLGQVFDSTAMKDKLLRIK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRVKIA KDGSGVINGWTQATDSILVFEVEVRVKIA KDGSGVSRDSAAPSPIPGADNLSADPV VSPFTSIFPKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMJPLIPPYQLDSNDSEI EDVNTLQGGGVPVTTSVYPSLQPAHFALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSAA VSGNAQSFQPYAGMQAYPQASAVTSQLQ PVRPLYPAPLSQPPHPQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMMSNIQRIIGENER LKQEILEKSNRIEGONDKISELIERNQYVEGS NILMMEKRNSLQTATESTMMSNIQRIIGENER LKQEILEKSNRIEGONDKISELIERNQYVEGS NILMMEKRNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKEFIEL QMQLTESLKETDLIRGQTKVQAALLSQCAEEIDEI RESEYLEKRILSTRKKRSQERSQAEEIDEI RESEYLEKRILSTRKKRSQCBERQAEEIDEI RESEYLEKKINGSVGVGRAAS DEPSEKVKKRMOQYPGJSRRFEELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRFSQQSASASSGQPALAS DEPSEKVKKRMOQYPGJSRRFEELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRFSQQSASASSGQPALNTERP ESPMYPSEQVVEAVPLOPLTTSQDGHRR KGDSBAEALSEIKDGSLPPELSCPSHVLGFP DEPSELVSKLKMOQYPGJSRLSFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRFSQQSASASSGQPALNTERP ESPMYPSEQVVEAVPLOPLTTSQDGHRR KGDSBAEALSEIKDGSLPPELSCPSHVLGFP SEPMYPSEQVVEAVPLOPLTTSQDGHRR KGDSBAEALSEIKDGSLPPELSCPSHVLGFP SEPMYPSELGONSMASRSGGDEEDDEVSINKGR PPPTPLFGDDDDDDDDDIDDUGG	1 .	1					
NSALIMVVVGSLSQYQLMKHQSRGFYOKQQ YLRWPEDLQDVGWFLHKLSQQILKEKKEKK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGVRPAKQHAAAAFYDVGGDRPWDS GNTQLPPRIPVKANAMFGAGDEDTDTLSPS GGARLASLFGLDQAAAGHGMFFQYTAPKQP KKGQGTAATGNQATEKTAPATMSTPTLVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPVTVARHVNTELMVRPNNY STFYDDQRQNWSIMFSEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGMPQNHVLGYBSTANKDKLIRIK LGSGKVTKGWEDGMLGMKKGGRELLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSSPTSIPFKSGEPALRTKSNSLSQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTPSVOPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHDSA VSGNAQSFQPYAGMQAYAYPQASAAVTSQLQ PVRPLYPAPLSQPFHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEDNIKSELIERNGRYVEGS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHQLKMTHAKKETEL QQMQLTESLKETDLLERGQLTKVQAKLSELQET SEQAQSKFKSEKQNRQQLEKVTSLEEELTDL RYSEGELENLISERKKSAQERSQAEEDDEI RKSYQEELDKLRQLLKKTRVSTDOAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELENKSSQMSGVAAASS DPSEKVKKIMNQVPGSLREFFELESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQASASSGQPQAPLNREPP ESPMVPSEQVEAVPLPPQALTTSQDGHRR KGDSEAEALSEKDGSLPFELESYNGRTI LGTIMNTIKMVTLQLLNQQECKEESSSEEEE EKAEERPRRPSQEQASASSGQPQAPLNREPP ESPMVPSEQVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEKDGSLPFELESTSRRVLGFP TSIPPEPLGFVSMDSECEESLAASPMAAKPDN PSGKVCVEVAPDGPLQESSTRISLTSDPEE GDPLALGPESPGPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPFPTLEFGDDDDDDDDURG	i	1	ł		}		
SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL SIGNAL Signal S					j		
KDNNIPLQTVATIS	Į.		į.				
816 2166 A 6646 I 3811 RDRAGVRPAGKQHAAAAYDVGGDRPWDS GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGHGNEFTQYTAPKQP KKGQCTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQVYKQGKFGAAVLGNHTTK EYRILLYISQQQPVTVARIHVNFELMVRPNNY STFYDDQRQNWSIINFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLILK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEFALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMJPLPPQLDSNDSEI EDVNTLQGGGQPVYTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSRQPYAGMQAYAPQASAVTSQLQ PVRPLYPAPLSQPPHPQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHMTKVEELQKH SAGNSMLPSMSVTMETSMIMSNIGNIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NILMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAPKGKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKPKSEKJNKQLELKVTSLEELTDI RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEACHLLASAKGPELAQAA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQGQEKEESSSEEE EKAERPRRPSGCAPAELALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQGQEKEESSSEEE EKAERPRRPSGQVEESSLSAASSMGQAEPLINGERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPDVBMDSECEESLAASFMAAKVPDN PSGKVCVRRVAPDGPLQESSTRLSLTSUPEE GDPLALGPESSFEEDPGLAKEBVTSSTFPKIK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKARPKAQPEEDDEDVSMKGR PPPTTLPFGDDDDDDIDUNGG			i				
GNTQLPPRNPVKANAMFGAGDEDDTDLSPS GGARLASLFGLDQAAAGHGNEFPOYTAPKOG KKQQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPVTVARIHVNFELMYPNNY STPYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKOKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTDTSSLVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSSPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMIPLLPPQLDSNDSEL EDVNTLQGGQPPVTPSVQPSLQPAIPALPQ MTSQAPQPSVTGLQASALMQVSSLDSHSA VSGNAQSSQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTERMAVSKVADKMDHLMTKVEELQKF SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQELLEKSNRIEEQNDKISELIERNORYVEQS NLMMERRNNSLQTATENTQARVLHAQEKKA KVTELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRQCLTKVQAKLSELQET SEQAQSKFKSEKQNSKQLELKVTSLEEELTDL RVSKYGELDKLRQLKKTVSLEEELTDL RVSKYGELDKLRQLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQGESVCFACLALQA QITALTKQNEQHIKELEKNRSQNSGVFAAAS DPSEKVKKIMQQYPQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEKSVCFACLALQA QITALTKQNEQHIKELEKNRSQNSGVFAAAS DPSEKVKKIMQQYPQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEGKEESSESEEE EKAEERRRPSGQSASASSGQPQAPLNRERP ESSMYPSEQVVEEAVPLPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPPLSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKVPDN PSGKVCVREVAPDGPLQEESSTRLSLTSUPEE GDPLALGFESFGEPPPQLKKDDVTSTGFHKK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEDDEDEVSMKGR PPPTTPLFGDDDDDDIDWLG	916	2166		6646	1	2011	
GGARLASLFGLDQAAAGHGNEFFQYTAPKOP KKGQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPVTVARHIVNFELMVRPNNY STFYDDQRQNWSIMFSESKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGERARIKKSNSLSEQLAINTSSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQYYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPHHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTMETSIMISNIQRIQENER LKQEILEKSNRIEGEQNDKISELIERNQRYVEQS NLMMEKRINSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVSKSCELDKLSRNKEASAGERSQAEEEIDEI RKSYGEELDKLRGLLKKTRVSTDQAAAEQLS LVQAELQTTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNRSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMYTLQLLKQDGCEKEESSSEEEE EKAEERPRPSQEQSASASSQQPQAPLNRERP ESSMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSSAEALSEIKDGSLPPGLSCTSPIRVLGFP TSIPPEPLGPVSMDSCEEESLAASPMAAKVPDN PSGKNVCVREVAPDGPLQESSTBLSLTSUPEE GDPLALGFESSFEEEE EKAEERPRPSQEQSSTBLSLTSUPEE GDPLALGFESSFEEEE GDPLALGFESSFEEEE EKAEERPRPSQEQSSTBLSLTSUPEE GDPLALGFESSFEEDEISTRUSCTSUPEE	010	2100	^	0040	*	3011	
KKGQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQPVTVARIHVNFELMVRPNNY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTOATDGILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLANTSPD AVKAKLISRMAKMGQPMLPLPPQLDSNDSEI EDVNTLQGGGQPVVTPSVOPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPHIFGGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNISQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAVGKETEL QMQLTESLKETULIRGQLITKVQAKLSELOE SEQAQSKFKSEKQNRKQLELKVTSLEEELITDL RVEKESLEKNLSERKKRSQERSQAEEIDEI RKSYQEELDKIRQLIKKTTVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRADYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHELEKNKSQMSGVEAAAS DPSEKVKKIMNQVPQSLREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAELSECIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKVPDN PSGKWCVREVAPDGPLQESSTRLSLTSUPEE GDPLALGPESPGEPOPPQLKKDDVTSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEEVSNKGR PPPTPLFGDDDDDDDDDDDWSLG		İ					,
AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQPVTVARIHVNFELMVRPNNY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDGLIVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPFGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGVYVTPSVOPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQMSTLMTEAR QHINTERMAVSKVADKMDHLMTKVEELQKH SAGINSMLPSMSVTMETSMIMSNIQRIIQENER LKQELLEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNKQLELKVTSLEELTDL RVEKESLEKNLSREKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQKVQLULKKTSVTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQKVQLULKKTSVTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYGVLVQLLGKSTSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVPGSLREFELEESYNGRTI LGTIMNTIKMVTLQLLNQCGCKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSSAEALSSEGEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSSAEALSSEGEEPELEESYNGRTI LGTIMNTIKMVTLQLLNQCGCKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSSAEALSSEGSEEPELEESYNGRTI LGTIMNTIKMVTLQLLNQCGCKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSSAEALSSEGSEPELEESYNGRTI LGTIMNTIKMVTLQLLNQCGCKEESSTEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSCAEALSSEGSEPEEPEDFOYMAKVPDN PSGKVCVCREVAPDGPLQESSTRLSLTSUPEE GDPLALGPESPGEPOPPQLKKDDVTSSTGFHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPSHHSQRSSLSGDEE DELFKGATLKALRPSHHSQRSSLSGDEE DELFKGATLKALRPSHHSQRSSLSGDEE	İ		l		į	i	
EYRILLYISQQQPYTVARIHVNFELMVRPNNY STFYDDQRONWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSGGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLANTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNISMLPSMSVTMETSMMSNIQRIQENER LKQELEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTUQLLNQQEQEKEESSSEEEE EKAEEPPRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPUSNDSECEESLAASPMAAKPDN PSGKVCVREVAPDGPLQESSTLSLTSUPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR			i				· · · · · · · · · · · · · · · · · · ·
STFYDDQRÖNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NILMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNKRQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNOVFGSLREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRPSQEQSASASSGOPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCPSHRVLGFP TSIPPEPLGPVSMDSECESLAASPMAAKPDN PSGKVCVREVAPDGPLQESSTRLSLTSUDEEE GDPLALGPESSGEFQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHASQRSSLSGDEE GDPLALGPESSGEFQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHASQRSSLSGDEE	I	ĺ			ĺ	ĺ	
KCNSTSSLDAVI.SQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGOPMLPILPPQLDSNDSEI EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTIMETSMIMSNIQRIIQENER LKQEILEKSNRIEEGNDKISELIERNQRYVEGS NILMMEKRNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELGTVOMEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMOVQYGSLRREFELEESYNGRTI LGTIMNTIKMVVLQLLQRQEKEESSSEEEE EKAEERPRRPSQEGSASASSGQPQAFLNRERP ESPMVPSEGVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKVPDN PSGKVCVREVAPDGPLQESSTRLSLTSDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSKAQPSEEDEEVSMKGR PPPTPLFGGDDDDDDDDWLG							
VAYTGWLFQNHVIGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAYGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMPLIPPQLDSNDSEI EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPHIFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQCTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLESRKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELOTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAERPRRPSQEQSASASSGQQAPLNERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSCEESSLAASPMAAKPDN PSGKVVCVREVAPDGPLQESSTRLSLTSUPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTFLEIGDDDDDDDDDDDDDDVLG	1	!					
LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPHFPGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYYEQS NLMMERNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSMSGVEAAAS DPSEKVKIMNQVPGJSREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAERPRRPSQEQSASASSGPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMMSECEESLAASPMAAKPDN PSGKVCVREVAPDGPLQESSTRLSLTSUPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTFLEIGDDDDDDDIDIWLG		1					-
CAVGSEGVIGWTQATDSILVFEVENRVKIA KDSGSDGHSVSSRDSAAPSPIFGADNISADPV VSPPTSPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTPSVQPSLQPAHFALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHINTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNINSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTSLKETDLLRQQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESSMVYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSSCEESLAASPMAAKYPDN PSGKVCVREVAPDGPLQESSTRLSLTSDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR	1						
KDSGSDGHSVSSRDSAAPSPFGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPLPPQLDSNDSEI EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PYRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRQQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVAQARDAYQQKLVQLQEKSVCFAKCLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKYPDN PSGKWCVREVAPDGPLQESSTRLSLTSDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSKHSQRSSLSGDEE DDELFKGATLKALPKAQPEEEDEDEVSMKGR	1		i				
VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVTPSVQSLQPAHFALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFAICLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAERPRRPSQEOSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEKKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKPDN PSGKVCVREVAPDGPLQESSTRLSLTSDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR			ļ				
AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVYTPSVQFSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMQVPQSLREEFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKPDN PSGKVCVREVAPDGPLQESSTRLSLTSDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR			İ				
EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKIPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDDDDVLG		}					
MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKIPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDVLG		l	i				
VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKONRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG							
PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLITESLKETDLLRGQLTKVQAKLSELGET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKVPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD					İ	:	
QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NILMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRQQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFAICLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSEEEE EKAEERPRRPSQEQSASASSQQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKVPDN PSGKVCVREVAPDGPLQESSTRLSLTSDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDDEVSMKGR PPPTPLFGDDDDDDDDIDWLG		İ			ļ		
SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKONRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKVPDN PSGKVCVREVAPDGPLQESSTRLSLTSVDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDIDWLG	1		J		}	ł	
LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFAICLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKVPDN PSGKIVCVREVAPDGPLQESSTRLSLTSUPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG							
NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKYPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDIDWLG	i				1	1	1 KUBII EKGNDIEEUVIURIGEI IEDVIUDAMEUG
KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDWLG	1			l	l		
QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG			1	Į	ł.	l	
SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG	İ		ſ		1	[
RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG	1		l	1			
RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG	1		1	1]	PACAGO EN EL CEDANAGY UEDGUY ECCIDE.
LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG			1		1	1	DECAUSE UKI DULI KALDAGERIAN YANG C
EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG	1	1	ł	1	1		I AUVEI ULUMEYRUBILI Y 6YAUBIL UUAU
QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG							
DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALITSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG		}	1		1	1	
LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG				1		!	
EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG]	ļ	}	}	1	1	L CTR OTTENANT OF FIGURE COORDER
ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG			1	1		Į.	EN VEED BE BECKER OF THE TOTAL
KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG			1		1	1	
TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG				1	}		ESTMYTSEQVYEEAVPLITQALTTSQUUHRR
PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG				1		1	
GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDWLG			ł		i	1	
ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDWLG							
DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDWLG			1				GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK
PPPTPLFGDDDDDDDDWLG .							
			j	1	J	1	
817 2167 A 6649 63 1073 FFRSSSDNGSPIROVE/HSTDAUGGDVMGI EG		L	L		<u> </u>		
OI ZIO A OUT O III	817	2167	Α	6649	63	1073	FFRSSSDNGSPIRQYE/HSTPAHQGPVMGLEG

	CEC IN	1/	CEC	Daniel de 3	Dec di - 4 - 3 ·	
	SEQ ID	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	NO: of peptide	поа	in NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
			USSN	location		I=Isoleucine, K=Lysine, L=Leucine,
	seq-		09/496	correspondi	corresponding to last amino	
	uence		914	•	acid residue	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first		Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						KS/ARNSQLRIVLVGKTGAGKSATGNSILGRK
i						VFHSGTAAKSITKKCEKRSSSWKETELVVVD
						TPGIFDTEVPNAETSKEIIRCILLTSPGPHALLL
						VVPLGRYTEEEHKATEKILKMFGERARSFMIL
]						IFTRKDDLGDTNLHDYLREAPEDIQDLMDIFG
						DRYCALNNKATGAEQEAQRAQLLGLIQRVV
						RENKEGCYTNRMYQRAEEEIQKQTQAMQEL
						HRVELEREKARIREEYEEKIRKLEDKVEQEKR
i)						KKQMEKKLAEQEAHYAVRQQRARTEVESKD
						GILELIMTALQIASFILLRLFAED
818	2168	A	6660	357	1890	APSGSWTRVVLTLDPCSLRSRSPRSLLDPGMP
						GISARGLSHEGRKQLAVNLTRVLALYRSILDA
			1	·	}	YIIEFF\TDNLWDTLPCSWQEALDGLKPPQLA
						TMLLGMPGEGEVVRYRSVWPLTLLALKSTA
]			j			CALAFTRMPGFQTPSEFLENPSQSSRLTAPFR
						KHVRPKKQHEIRRLGELVKKLSDFT/GLHPGC
						RRGLRPG\HLSRFMALGLGLMVKSIEGDQRL
						VERAQRLDQELLQALEKEEKRNPQVVQTSPR
1 1			}			HSPHHVVRWVDPTALCEELLLPLENPCQGRA
1						RLLLTGLHACG\DLSVALLRHFSCCPEVVALA
				1		SVGCCYMKLSDPGGYPLSQWVAGLPGYELP
						YRLREGACHALEEYAERLQKAGPGLRTHCY
1 1	_		-	ĺ	}	RAALETVIRRARPELRRPGVQGIPRVHELKIEE
į į	·					YVQRGLQRVGLDPQLPLNLAALQAHLAQEN
						RVVAFFSLALLLAPLVETLILLDRLLYLQEQA
l i			ì	·		LSP\GFHAELLPIFSPELSPRNLVLVATKMPLG
						QALSVLETEDS
819	2169	A	6661	65	2686	SGSGHCLAEAASMGPWGWKLRWTVALLLA
						AAGTAVGDRCERNEFQCQDGKCISYKWVCD
		ļ	}	ļ		GSAECQDGSDESQETCLSVTCKSGDFSCGGR
			ļ	!		VNRCIPQFWRCDGQVDCDNGSDEQGCPPKTC
}			1			SQDEFRCHDGKCISRQFVCDSDRDCLDGSDE
1						ASCPVLTCGPASFQCNSSTCIPQLWACDNDPD
1		1	l	1	ļ	
		í				CEDGSDEWPQRCRGLYVFQGDSSPCSAFEFH
			!			CLSGECIHSSWRCDGGPDCKDKSDEENCAVA
, 1						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC
				<u> </u>		CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM
	·					CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENILLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL
						CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP
	-					CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLARVDMRSCLTEGVEAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG/RGN/EKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED
820	2170		6656		4146	CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAGRGNEKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA
820	2170	A	6666	. 17	4146	CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL PIVVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA
820	2170	A	6666	. 17	4146	CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA LSLWPTSGEICGPGIDIRNDYQQLKRLENCTVI
820	2170	A	6666	. 17	4146	CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLAR\DMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAG\RGN\EKKPSSVRALSIVL PIVVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA

SEQ ID No. of nort peptide contide graphic permission of peptide solide graphic permission of peptide solide graphic permission of the permission of the permission of the permission of the permission of the permission of the permission of the permission of the permission of the permission of the permission of			r				
nucleotide sequence USSN (9496) location corresponding of the properties and the properties of peptide residue of peptide residue of peptide sequence per la control of the per	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence ### Seque	NO: of		hod	ID NO:		nucleotide	
sequence 914 914 915 916 916 917 917 918 918 918 918 919 919	nucl-	peptide	ł		nucleotide	location	
sequence 914 914 915 916 916 917 917 918 918 918 918 919 919	eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
uence Part Region Regio	seq-			09/496	correspondi		
mino add residue of peptide requence propried sequence propried sequence propried sequence propried sequence propried sequence provided se			١.	914		h	
residue of popidid sequence			· .		. –		
Peptide				1			
	1		ł	1		, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	
EMTNIKDIGI, YNLKNITRGARREKNADLCYL STYDWSLILDAVSNIYUONICPKEGGDLC GTMERKPMCEKTTINNEYNYKOWTYTREGOLC GTMERKPMCEKTTINNEYNYKOWTYTREGOLC GTMERKPMCEKTTINNEYNYKOWTYTREGOL MCPSTCKRACTENNEGCHPECLGSCSAPDN DTACVACRHYYYAGVCVRACPPNTYREGOW RCVDRDFCANILSASESDSGEVHIDGEMGEMGE CPSGFIRNGSOSMYCTPCGGPCFK VCEEEKT XTIDSVTSAQMLGGCTFKGMLLINREGNNIA SELNPMGLIEVYTGYVKIRRISIALVSLSTIK N.R.I.I.GEQLEGNSYFYVLDIONALOU, WD WDHRNLTIKAGKMYPAPRIPLCVSETYRME VTGTKGROSKGDDTRINNGERASCESDVLHF TSTTTSKNPEILTWSRPSPLDLASS SNSSQLIVKWNPSLPNONLSYTVRWORD DOROTYPRINCESSEDVLHF TSTTTSKNPEILTWSRPSPLDVSAS NSSSQLIVKWNPSLPNONLSYTVRWORD DOROTYPRINCESSEDVLHF TSTTTSKNPEILTWSRPSPLDVSAS NSSSQLIVKWNPSLPNONLSYTVRWORD DOROTYPRINCESSEDVLHF TSTTTSKNPEILTWSRPSPLDVSAS NSSSQLIVKWNPSLPNONLSYTVRWORD DOROTYPRINCESSEDVLHF TSTTTSKNPEILTWSRPSPLDVSAS NSSSQLIVKWNPSLPNONLSYTVRWORD DOROTYPRINCESSEDVLHF TSTTTSKNPEILTWSRPSPLDVSAS NSSSQLIVKWNPSLPNONLSYTVRWORD DOROTYPRINCESSEDVLHF TSTTTSKNPEILTWSRPSPLDVSAS NSSSQLIVKWNPSLPNONLSYTVRWORD DOROTYPRINCESSEDVLHF TSTTTSKNPEILTWSRPSPLDVSAS NSSSQLIVKWNPSLPNONLSYTVRWORD DOROTYPRINCESSEDVLHF TSTTTSKNPEILTWSRPSPLDVSAS NSSSQLIVKWNPSLPNONLSYTVRWORD DOROTYPRINCESSEDVLHF TSTTTSKNPEILTWSREENTYMGVA NTTMSSSSSNATTAADTYNTERGENTYMGVA NTTMSSSSSNATTAADTYNTLEGPTOPHTYME REPRINCESSEDVLHFT TSTTTSKNPEILTWSREENTYMGVA NTTMSSSSSNATTAADTYNTLEGPTOPHTYME REPRINCESSEDVLHFT TSTTTSKNPEILTTSKEENTYMGVA NTTMSSSSSNATTAADTYNTLEGPTOPHTYME REPRINCESSEDVLHFT TSTTTSSTATAG NTTMSSSSSADATAGORYTHOMATAGORYTH ARTYPEACAGAM NATIONAL STATAGORYTH ARTYPEACAGAM NATIONAL STATAGORYTH ARTYPEACAGAM NATIONAL STATAGORYTH ARTYPEACAGAM NATIONAL STATAGORYTH ARTYPEACAGAM NATIONAL STATAGORYTH NATIONAL ST			1	l			
STVDWSLILDAVSNNTVIVGNIKPREGGDLC	<u> </u>		 	 	sequence		
GTMEEKPMCEKTTINNEYNYRGWTYNRCGK MCPSTCKRACTENBECCHPECLGSCAPDN DTACVACRHYTYAGOVCYACCPRITYREGW RCVDRDFCANIL SAESSSEGFWIEDGECMGE GPSGFIRNGSQSMYCTPECGPCFKVCEEEKKT KTIDSVTSAQMLQGCTIKGGLLNRGKOLNIRGKOLNI BEIRDMGLIEVTOVVKIRHSHALVSLSFLK MIRLILGERQLEGNYSFYVLINONLQQLWD WHERHLITLAGKMYRAPPRICVELSFYRMEE VTGTKGRQSKGDNTRNNEGRASCESDVLHF 1STITTSKNIIITWHRYPDYBLDISTYTYK EAPPKNYTEYDGQDACGSNSWNNVDVDULPP NRDWPOFILLHGLKWYRPDYBCILISTYTYK EAPPKNYTEYDGQDACGSNSWNNVDVDULPP NRDWPOFILLHGLKWYPDYBCHLSTYTYK EAPPKNYTENGLLHGLKWYRONGO QDGYLYRHNYCSKDKIPRKYADGTIDIEEVT FENFETVCGGEKGPCACCKFKTEAKAQEKEE ASYRKVFENI-HINSIFVRPERRRRDVMQVA NTIMSSRNTTAADTNIDPEELTTYPF ESRVDNKERTVISNLRPFTLYRDHISCHHAE LIGGSASNFVARTMPAGADDIPGFYTWEP RESNIFLKWFEENINGLLMYEKYGSQVE RESVENGEYSKYGGAKLNRLNPGNYTARI QATSLSGNGSWTDPYFFYVQAKYVKNPHILI ALPIVAVLLIVGGLUMLYVHRKRNSRLGN GVLYASSVAPPSSAADVVVPDEWEVAREKIT MSRELGOGSFGMYVGAKGAKLNRLNPGNYTARI QATSLSGNGSWTDPYFFYVQAKRYKENHHILI ALPIVAVLLIVGGLUMLYVHRKRNSRLGN GVLYASSVAPPSSAADVVVPDEWEVAREKIT MSRELGOGSFGMYVGAKGVKLOPEFTRV ARTVNEAASMEREFTINEASYMKEFNCHH VYRLLGVYSOGOPTLVTMELMTRGDLKSYLR SLRFPENNPYLAPPSLSKMJOMAGELDOM AVLNANIEPSTSAMGONAGELADOM AVLNANIEPSTSAMGONAGELADOM AVLNANIEPSTSCHAKGKGKLUW MKMCWQYNFKMRPSELLSKKEEMBEFGTRE KGOVETTYSDWSGVVLWELATLAGEPYO GLSNEQVLEVVMSGGCLLDKPDNCTPOMLFEL MKMCWQYNFKMRPSELLSKKEEMBEFGTRE SESTENDAMGNAGALADOMAGELADOWSTELL BEZI ZITI A 6691 106 825 GRVLFRGCGVGHKGQVLMGTTELAGDPYO GLSNEQVLEVVMSGGCLLDKPDNCTPOMLFEL MKMCWQYNFKMRPSELLLSREEMSKKE SEX DTSNERVERGEMGIGHERVYT VISKARCRNTILARRKGHMGIGKKGTAN ARMPEKVTWMRRMILIRALLRRYRESKEYR ERKIDNAMSRCGLAKTENSKRETKREAR REERLOAKKEULADOAGARSKYKREAR REERLOAKKEULADOAGARSKYKWEA UDPRETNEIANANSRQQIKKLIKDGLIRKPYT VISKARCRNTILARRKGHMGIGKKKGTAN ARMPEKVTWMRRMGILRRLLRRYRESKEYR ESKIDRHMYHSI, VKKONVYKKRKRILMEH HKIKLADAKARKILLADQAGARRSKYKREAR REERLOAKKEULADOAGARSKYKNEAR REERLOAKKEULTHERVERSKYRY SCHIKDRAMSTSNASSPAGSPLETTINGROCHTELKPCC SKAKANAGKREGGEGULDTILPKPCC SKAKANAGKREEGGGFEULISTICEWYTENSSTNPAA SSSALSSSSYPVSSSAAROUSGISTISSSSSTNPA			1	l			L TELEVISION CONTRACTOR CONTRACTO
MCPSTCGERACTENNECCHPECLGSCSAPDN DTACVACRHYTYAGVCYNACPTYREGW RCVDRDFCANILSAESSDSGFVHDGEGCNG CPGGFRNGSGSWCYDCGCGPCPK VCEEKKT KTIDSVTSAQMLQGCTIFKGNLLINIRGGNNA SELENFMGLIEVVTGVVKIRHSHVSLSFLK NRLLLGEGLEGNYSFYVLDNONLQQLWD WDHENLTIKAGKMYFADPRILAVSLSFLK NRLLLGEGLEGNYSFYVLDNONLQQLWD WDHENLTIKAGKMYFADPRILAVSLSFLK NRLLLGEGLEGNYSFYVLDNONLQQLWD WDHENLTIKAGKMYFADPRILAVSLSFLK EAPFENTTEYDGQDACGSNSWANVDVDLYF FISTITISENRIILWH PXPPDYBRDLISFTVYFK EAPFENTTEYDGQDACGSNSWANVDVDLYJSAS NSSSQLIVKWNPPSLPMGNLSYTIVRWQRQP QDOYLYRINYCSKDKIPRKYADDIBEVT EMPKTEVCGGEKGPCACPKTEAEKQAEKGE AFYRKYFENHHNSFVPFREKREDVMQVA NTIMSSRSRNITAADTYNTDFFELETEYPFF ESRVDNKERTVISNLRPFTLYRBIDISCHHEAE KLGCSASNFVFARTMFABGADDIPGFVTWP RPENSIFLKWPEPEPMPGLLMYEIKYGSQVE DQRECVSRGEYRKYGGAKLNRLNPGNYTAGH QATSLSGNGSWTPVFTVQAKKVENFHILLI ALPVAVLLIVGGLVIMLYVFHRKRNNSRLGN GVLYASVSPEYFSAADVVYPDEWVAREKITI MSRELGQGSFGMVYEGVAKGVVXGDFETW AKTVRAASMREREITLNESSVAKKEPNCHH VVRLLGVVSQQPTLVIMEIMTRGDLKSYL SLRPEMENNPULAPPSISKMIMGMAGEJADM AYLMANKFVHRDLAARNCMVAEDFTVKIG GSNGMYPFWKGGKLGLPVRWMSPESL KDGVFTTYSDWSFGVVLWEIATTAEQDLYST SLRPEMENNPULAPPSISKMIMGMAGEJADM AYLMANKFVHRDLAARNCMVAEDFTVKIG FGMTBDIYETDYYPKGGKGLLPVRWMSPESL KDGVFTTYSDWSFGVVLWEIATTAEQDLYST GGNGMYPFWKGGGLLBKPDNCPDMLFEL MRMCWQYNPKMRFSFLEISSIKEBMEPGFRE VSFYYSEENKLPPEPELDLEPENMCPLMYBMSPLDPS ASSSLPLPDRIGGKAENGRAPLPGSSTC GRVLFRGGGVGHKGQVLWEIATTAEQPYS SINGVCNSSMLRILAKGNINGRILAKENGLMEH EIKLKADKARKKLLADQAEARSKYTKEARK REERLQAKKEEIKITLSKEGETKK SCHCHAMANARCHARKGRINGRIKKGTAN ARMPEKVTWARRARRIERLRSKRYRESKAYR ESKKEDHMYTSLYLKWKGNVFKRILMEH EIKLKADKARKKLLADQAEARSKYTKEARK REERLQAKKEEIKITLSKEGETKK SCHCHAMANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARCHARTHERPRESCOCT SCHKKRANARC			1	j		i	
DTACVACERTYYAGVCYPACPENTREEGE RCVDRDFCANILASESSDSEGVHIDGECMQE CPSGFIRNGSQSMYCIPCEGPCPK VCEEEKKI KTIDSVTSAQMLQGCTIFK KOHLINIRGRONIA SELENFMGLIEVYTGYVKIRHSHAL VSLSTY NIRLI LGEPOLEONYSFYLJONOL QQL VIV WDIRNLTIKAGKMYFAFNYKLCVSETYRMEE VTGTTKGRQSKGDINTRNNGERASCESDVLH TSTITTSKNRIITI WHEYRPPY XRDLISITYVYK EAPFKINVTEVDGQDACGSNSWAMAVDVDLY NKDVEFGILLHGLKPWTQYAVYVKAVTLTM VENDHIRGAKSELI YIRTNASVPSIPLDVISAS NSSSQLIVKWNPPSLPKMOLSYTYVRWGRQP QDCYLYRINYCSKDKIPRKYADGTIDIEEV EPIKKTEVCGGEKGPCCACPKTEAKQAEKEE AEYRKVFENI-HINSIFVPRPERKRDVMQVA NITIMSSRSNITAADTYNTIDPELEITE YPFF ESRVDNKERTVISINLRPFTLYRIDHISCNHEAE KLGCSASNFVFARTMPAGADDIEVTWEP RPENSIFLKWPEPENPIGLILMYEIKYGSQVE DRECVSRQEYRKYGGAKLNRIHONYTARI QATSLSGNGSWTDPVFFYVQAKRYENPHILI ALPVAVLLIVGGLVMLVYHREKRNSKLIGN GVLYASVNPEYFSAADVYYPDEWEVAREKIT MSRELGGGSFGMVYEGVAKGVVEGVAKGVEPETRV AIKTNNEAASMRERIEFINEASVMKEFNCHH VVEILLGVVSQGGPUTLVIMELMRKRNNSKLIGN GVLYASVNPEYFSAADVYYPDEWEVAREKIT MSRELGGGSFGMVYEGVAKGVVEGVAKGVVEGVAKGVVEGVAKGVVIGV GLSNEGVIRVYVERGGAKLIRATGNSKLIGN GVLYASVNPEYFSAADVYPDEWEVAREKIT MSRELGGGSFGMVYEGVAKGVVICHATATAGPGVT GLSNEGVIRVYVERGGAKLIRATGNSKLIGN GVLYASVNPEYFSAADVYPDEWEVAREKIT MSRELGGGSFGMVYEGVAKGVVEGVAKGVVEGVAKGVVICHATATAGPGVT AIKTNNEAASMRERIEFINEASVMKEFNCHH VVEILLGVVSQGGPUTLVIMELMREKYNSKLIGN GVLYASVNPEYFSAADVYPDEWEVAREKIT MSRELGGGSFGMVYEGVAKGVVICHATATAGPGVT GLSNEGVIRVYVEGGALLARATAGPGVT GLSNEGVIRVYVEGGALLARATAGPGVT GLSNEGVIRVYVEGGALLARATAGPGSTACH VVEILLGVVSQGGACHAGVILWEIATLARATAGPGVT GLSNEGVIRVYVEGGALLARATAGPGSTACH VVEILLGVVSQGGACHAGVILWEIATLARATAGPG GSCGVGHYKGGVLWEIATLARATAGPGVT GLSNEGVIRVYVEGGALLARATAGPGSTACH VSTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG				1			
RCVDRDFCANILSAESSDSGFVHHDGECMCK CRSGFRNGSGSWYCPCEGCPCEBEKT KTIDSVTSAQMLQGCTIFKGNLLINIRGRONIA SELENFMGLIEVVTQVKIRHSHALVSLSFLK NERLILGEQLEGNYSFYVLDNONLQQLWD WDIRNLTIKAGKMYFANPKLCVSEIYRMEE VTGTKGRQSKGDINTRNNGERASCESDVLHF TSTTTISKRITITHHRYRFDYRDLISFITYVY EAPFKNVTEVDGQDACGSNSWHMVDVDLPP NKDVEPGILHGLKPWTQYAVYVKAVTLTI VENDHIRGAKSELTYIRTMSVSPSIPLOVLSAS NSSSQLIVKWNPPSLPNGNLSYYTVRWQRQD QDGYJYRHNYCSKOKIPIRKYADGTIDIEEVY ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE AFYKRVFENHINSIPPTPREBUBLISTYYYGV QDGYJYRHNYCSKOKIPIRKYADGTIDIEEVY ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE AFYRKVFENHINSIPPTPREBUBLISTOR AVKRYFENHINSIPPTPREBUBLISTOR REPRINTENTER PROBLEMENT OF THE PROBLEMENT	1			1			MCPSTCGKRACTENNECCHPECLGSCSAPDN
CPSGFIRNGSOSMYCIPCEGEPKVCEEKKI KTIDSVTSAQMLQGCTIFKGNLINIRGGNNIA SELENFMGLIEVVTGYVKIRHSHALVSLSFI- NELL GEEDLESNYSFYLJENDNLQQLWD WDHRNLTIKAGKMYFAFNPKLCVSETYRMEE VTGTKGRQSKGDINTRNNGERASCESDVLHL TSTITISKNRIIT WHEYRPFYDRIDISTTYYFK EAPFKNYTEYDGQDAGGSNSWMWDVDUDL NKDVEPGILLHGLKPWTQYAVYVKAVTLTM VENDHIRGAKSELIYIRTNASVPSIPLDVLSAS NSSSQLIVKWNPPSEPKONLSYTYVRWQRQP QDGYLYRINYCSKOKIPIRKYADOTTDIEER FERVEVGEKGCPCACPKTEARKQAEKEE AEYRLYFENFLHNSIFVPRPERKRDVMQVA MITIMSSRSNITAADTYNITDIBELTETYPFF ESRVDNKERTVISNLRPFTLYRIDHISCNHEAE KLQCSASNFVFARTMPAGIADDIRGYTWR PRENSIFLKWPEPENPNGLLMYBERYGSQVE DQRECVSRQEYRKYGGAKLNRLNPGNYTABL QATSLSGNGSWTDVFFFVQAKKINRHHLII ALPVAVLLIVGGLVIMLYVFHRKRNINSRLOR GVLYASVNPEYSAADVYYPDEWVAREKIT MSRELGQGSFGMVYEGVAKGVXCDFEFTIV ALFANDAR SERVENSERSERSTANDSVANGEFKCHH VVRLIGVVSQQQPTLVIMEIMTRGDLKSYLR SLRPPEMSNPVLAPPSILSKINGMAELDGE AYLINANKFYHDDLARRICMVAEDFTVKIG GSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRNCWQNYNKRMSFSLEISISKEEMBFGIRE VSFYYSEENKLPEPELDLEPENMESVYLDGD FOMTRDIYETDYYRKGGKGLLPVRWMSPESL KDOVFTTYSDVWSFGVLWELATLABQPYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRNCWQNYNKRMFSFLEISISKEEMBFGRE VSFYYSEENKLPEPELDLEPENMESVYLDGD FOMTRDIYETDYRKGGKGLLPVRWMSPESL KDOVFTTYSDVWSFGVLWELATLABQPYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRNCWQNYNKRMFSFLEISISKEEMBFGRE VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLARGKKNTKAEK SESSENFINSKRYENGTCSCGLEETIK SGDNCNACVLLLKRWKKLPAGSKKNWNIN VDAARAPSLKLTKLKPKVKTNYKNRGLCHETH SGDNCNACVLLLKRWKKLPAGSKKNWNIN VDAARAPSLKTTK,KPKVKTNYFSFLEIDL TYWKRQKICCGNYKGRGEVLIDTHLFKPCC SNKKANAAKKREEQPPPLISTIQEWTEFYK VVPARAGSAAPPTGSSAAPGGSAAPTTSNSSSTTNSSSTTSSGFGS SSSASSTPYDSSAAAGGGATPG NAGPLAPNGSAAPPAGSAAPPTSNSSSTTSSGFGS	1	ļ .			i		DTACVACRHYYYAGVCVPACPPNTYRFEGW
CPSGFIRNGSOSMYCIPCEGEPKVCEEKKI KTIDSVTSAQMLQGCTIFKGNLINIRGGNNIA SELENFMGLIEVVTGYVKIRHSHALVSLSFI- NELL GEEDLESNYSFYLJENDNLQQLWD WDHRNLTIKAGKMYFAFNPKLCVSETYRMEE VTGTKGRQSKGDINTRNNGERASCESDVLHL TSTITISKNRIIT WHEYRPFYDRIDISTTYYFK EAPFKNYTEYDGQDAGGSNSWMWDVDUDL NKDVEPGILLHGLKPWTQYAVYVKAVTLTM VENDHIRGAKSELIYIRTNASVPSIPLDVLSAS NSSSQLIVKWNPPSEPKONLSYTYVRWQRQP QDGYLYRINYCSKOKIPIRKYADOTTDIEER FERVEVGEKGCPCACPKTEARKQAEKEE AEYRLYFENFLHNSIFVPRPERKRDVMQVA MITIMSSRSNITAADTYNITDIBELTETYPFF ESRVDNKERTVISNLRPFTLYRIDHISCNHEAE KLQCSASNFVFARTMPAGIADDIRGYTWR PRENSIFLKWPEPENPNGLLMYBERYGSQVE DQRECVSRQEYRKYGGAKLNRLNPGNYTABL QATSLSGNGSWTDVFFFVQAKKINRHHLII ALPVAVLLIVGGLVIMLYVFHRKRNINSRLOR GVLYASVNPEYSAADVYYPDEWVAREKIT MSRELGQGSFGMVYEGVAKGVXCDFEFTIV ALFANDAR SERVENSERSERSTANDSVANGEFKCHH VVRLIGVVSQQQPTLVIMEIMTRGDLKSYLR SLRPPEMSNPVLAPPSILSKINGMAELDGE AYLINANKFYHDDLARRICMVAEDFTVKIG GSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRNCWQNYNKRMSFSLEISISKEEMBFGIRE VSFYYSEENKLPEPELDLEPENMESVYLDGD FOMTRDIYETDYYRKGGKGLLPVRWMSPESL KDOVFTTYSDVWSFGVLWELATLABQPYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRNCWQNYNKRMFSFLEISISKEEMBFGRE VSFYYSEENKLPEPELDLEPENMESVYLDGD FOMTRDIYETDYRKGGKGLLPVRWMSPESL KDOVFTTYSDVWSFGVLWELATLABQPYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRNCWQNYNKRMFSFLEISISKEEMBFGRE VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLEPENMESVYLDGSTC VSFYYSEENKLPEPELDLARGKKNTKAEK SESSENFINSKRYENGTCSCGLEETIK SGDNCNACVLLLKRWKKLPAGSKKNWNIN VDAARAPSLKLTKLKPKVKTNYKNRGLCHETH SGDNCNACVLLLKRWKKLPAGSKKNWNIN VDAARAPSLKTTK,KPKVKTNYFSFLEIDL TYWKRQKICCGNYKGRGEVLIDTHLFKPCC SNKKANAAKKREEQPPPLISTIQEWTEFYK VVPARAGSAAPPTGSSAAPGGSAAPTTSNSSSTTNSSSTTSSGFGS SSSASSTPYDSSAAAGGGATPG NAGPLAPNGSAAPPAGSAAPPTSNSSSTTSSGFGS	l	ł	l	1			RCVDRDFCANILSAESSDSEGFVIHDGECMOE
KITIDSVTSAQMI QGCTIKGNILLINIRGONIA SELENFMGILEVTGYVKIRISHALVSLSPLK NLRLI.GEGQLEGNYSFYVLDNQNLQQUWD WDHRNITIK AGKMYFAFIPKLCVSEIYRMEE VTGTKGRQSKGDINTRNIGERASCESDVLHF TSITTSKNRIIITWHRYRPPDYRDLISFTVYYK EAPFRIVTEYDQQDACGSISWIMWDVDLPP NKDVERGILLHGLRPWTQYAVYVKAVLTIM VENDHIRGAKSELLYRINASVSPIPLUVLSAS NSSSQLIVKWNPFSLPINGRLSYYIVRWQROP QDGYLYRHNYSKSMEPREKYADGTIDIEEVT ENKITEVCGGEKGPCCACPKTEAEKQAKEE AFYRKYFENFILHISIFVPPFBERRRDWQVA NITIMSSRSRITIAADTYNITIDFELLETEYPF ESRVDNKERTVISNLRPFILYRIDHISCINHEAE KLGCSASNFVARTIMPAEGADDIPGPYWEP RPENSILLKWPEPENFINGLIKYRIGHISCNHEAE KLGCSASNFVARTIMPAEGADDIPGPYWEP RPENSILLKWPEPENFINGLIKYRIGHISCNHEAE KLGCSASNFVARTIMPAEGADDIPGPYWEP RPENSILKWPEPENFINGLIKHRIGHEKYGSQVE DQRECVSRQEYRKYGGAKLINILNIGNITTARI QATSLSGNGSWTDPVFFVVARFINSHTIN ALPVAVLLIVGGLVIMLYVFHRKRINSRLGN GVLYASVNEPYFSAADVYVPDEWEVAREKIT MSRELQGGSFGMVYEGVAKGVVKDEPETRV AIKTVNEAASMEREIFINEASVYMERNICHH VVKLLGVVSQGQFTLVMEHEMTRGDLKSVLR SLRPEMENNPVLAPPSLSKMQMAGEIADGM AYLNANKFYHRDLAARNCMAEDFTVKIGD FGMTRDIVETDYYRKGGKGLLPYRWMSPESL KDGYFTYSBUVSGVVLWEIATLARQPYQ GLSNEGVLRFVWBGGGLLDKPDNVCPDMLFEL MRRWQVYNFMMPSFSCLISKEMEPFGRE VSFYYSEENKLPEPELDLEPENMESVPLDPS ASSSLPLPDRINGHALARCHENSERSKYR SSSSLPLPDRINGHALARCHENSERKYR SSSSLPLPDRINGHALARCHENSERKYR SSSLPLPDRINGHALARCHENSERKYR SSKLPLPRINGHALARCHENSERKYR SKKUDRIMYHSLYLKVKGNYFKNKRILMEH BEKLKADKARKKLLADQAEARSKKKEARK REPERLQAKKERLADQAEARSKKKEARK REPERLQAKKERLADQAEARSKKKEARK REPERLAGKKERLINTLSKEREFTKK 822 2172 A 6715 772 21 DFRPGILLPRKKMFGFHKPKMYMSEGCKII SGDICMACVLLUKRWKLLPAGSKKNWMH VVDARAGPEKLTILLPRKYKILSGORGILFT SGDICMACVLLUKRWKLLPAGSKKNWMH VVDARAGPEKLTTLLPRKYKKLLPAGSKKNWMH VVDARAGPEKLTTLLPRKYKKLLPAGSKKNWMH VVDARAGPEKLTTLLPRKYKKLLPAGSKKNWMH VVDARAGPEKLTTLLPRKYKKLLSGORGILFT SGDICMACVLLUKRWKLLPAGSKRNWMH VVDARAGPEKLTTLLPRKYKKLLPAGSKKNWMH VVDARAGPEKLTTLLPRKYKKLLPAGSKKNWMH VVDARAGPERLYTTLLPRKYKKLNSGORGHETTINSSTRYAA SSSSAGSSSVPYSSSSSAGSVPTSNSSSSTRYA SSSSAGSSVPYSSSSAAGSAFTSNSSSSTRSAA SSSSAGSSVPYSSSSAAGSAFTSNSSSSTRSSGFGCG SKKAVAAAEKEEGGGFPLIPSTGEWVTEVFM AGGELLPRKAMFER SSSAGSSVPYSSSSAAGSAFTSNSSSSTRSSGFSC		1					
SELENFMGLIEVVTGYVERRISHALVSLSTLK NLRLIGEGOLEGNYSFYVLDNQLOLWD WDHRNLTIKAGKMYFAFNPKLCVSEIYMEE VTGTKGRQSKGDNTRNIGERASCESDVLHF TSTTTSKNRIIITWHRYRPPDYRDLISFTVYYK EAPFRAVTEYDOQDACGSNSWNMVDVDLPP NKDVEPGILLHGLKPWTQYAVYVKAVTLTM VENDHIRGAKSEILYRTNASVSEIPLJVLSAS NSSSQLIVKWHPPSLIVGRISLSYTYRWGRQP QDGYLYRHNYCSKDKIPIKYADGTIDEVT ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE AFYRKYENFELHNSIVYPRPERKRDVMQVA NITIMSSRSRNTADTYNTDIEVT ESRVDNKERTVISNLRPFTLYRDDHSCNHEAE KLGCSASNVFARTMAPBAGDIPGPYTWEP RPENSIFLKWPEPENPNGLILMYEIKYGSQVE DQRECVSRQEYRKYGGAKLNRLNFGNYTARI QATSLSGNGSWNTDVFFFYVQAKEYENFHLII ALPVAVLLIVGGLVIMLYVFHEKRNNSRLGN GVLYASVNPEYPSAADVVYPEWEVAREKTH WSRELGQGSFGMYYEGVAKGVKKDEPTEV AKTYNEAASMKREITEFILMEASVMKEPTCHHA SRELGQGSFGMYYEGVAKGVKADEPTEV AKTYNEAASMKREITEFILMEASVMKEPTCHHA VVRLLGVVSQGOPTLVIMELMTRGDLKSYLR SLRPPMENNPVLAPPSLSKMIQMAGGEADGM AYLNANKFVRIDLAARASVMKEPTCHHE VVRLLGVVSQGOPTLVIMELMTRGDLKSYLR SLRPPMENNPVLAPSLSKMIQMAGGEADGM AYLNANKFVRIDLAARASVMKEPTCHHE VVRLLGVVSQGOPTLVIMELMTRGDLKSYLR SLRPPMENNPVLAPSLSSKMIQMAGGEADGM AYLNANKFVRIDLAARASVMKEPTCHHE VVRLLGVVSQGOPTLVIMELMTRGDLKSYLR SLRPPMENNPVLAPSLSSKMIQMAGGEADGM AYLNANKFVRIDLAARASVMKEPTCHHE VVRLLGVVSQGOPTLVIMELMTRGDLKSYLR SLRPPMENNPVLAPSLSSKMIQMAGGEADGM AYLNANKFVRIDLAARASVMKEPTCHHE VVRLLGVVSQGOPTLVIMELMTRGDLKSYLR SLRPPMENNPVLAPSLSSKMIQMAGGEADGM AYLNANKFVRIDLAARASVMKEPTCHHE VVRLLGVVSQGOPTLVIMELMTRGDRWMSPSTL KDGVPTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVVMBFSGGLIBSISKEEMBEFGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPRRHSGGKAASSVLRCGKKKW UPPERTNEIAANASQQRKKGTAL PROMTHER VSRYSEENKLPEPEELDLEPENMESVFLDSP SSKLIDRHMYHSLVKLVKGNINGRICKSTL VHSRAACRKNILADQAEARSKTKEAAK REERLQAKKEELBKTLSKEEETKK REERLQAKKEELBKTLSKEEETKK SGDICNACVLLUKRWKKLPAGSKKINWNI VVDARAGPSLKTTLKFRKVKTLSGORIKST QISKLQKEFKRÜNSDAHSTTSSSSFOSGOT SGKKSSSSIKTPTSSNSTSSGFOSG SSKSSSSSSTPTSSSSFOSGOT SSKKAAAAAEREGGPPELPISTQEWVTEVFM TVNOFRWTGSDTOVGPFGOSTGCHETTR GGDICCAACVLLUKRWKKLPAGSKKINWNI VVDARAGPSLKTTLKFRKYKTLSGORIKST QISKLQKEFKRÜNSDAHSTTSSSSFOSGOT SSKKAAAAAEREGGPPELPISTQEWVTEVFM TVNOFRWTGSDTOVGPFGOSTGCHETTR SGDICCAACACCHETTEVTEVFM TVNOFRWTGSDTOVGPFG	1				ļ		•
NLRLILGEGLEGNYSFY/LDNQNLQQLWD WDMRNITIKAGKMYFAFIPKILCVSEIYRMEE VTGTKGRQSKGDINTRNIGERASCESDVLHF TSTITSKNRIIITWHRYRPPDYRDLISFTVYYK EAPFINVTEYDQDACGNSWNMVDVDLPP NRDVEPGILLHGILRPWTQYAVYVKAVITIM VENDHIRGAKSELIVIRINASYSPIPLUVLSAS NSSSQLIVKWNPSLFIPGILSFTYYRWQRQP QDGYLYRHNYCSDKUPIRKYAPGTIDIEVT ENRIFEVCGGEKGPCCACPKTEAEKQAEKE AFYRKYEPSI-HINSIEVPPPERISHRUMQRQP AFYRKYEPSI-HINSIEVPPPERISHRUM ANTIMASSRSRITTAADTYMITDPEELETEYPF ESRVDNKERTVISNLRPFTLYRIDHISCNHEAE KLGCSASNFVFARIMFAEQADDIPGPYTWEP RPENSIFLKWPEPENPIGILIMYEIKYGSQVE QDQECVSRQEYRKYGGAKINJRGMYTEVYHKRRINSRLGN GVLYASVNPEYSAADVVYPDEWEVAREKIT MSRELGQGSFGMYTECVAKVEPETRY AIKTYNEAASMERIFETNEASVMKEPKCHH VVRLIGVVNGQPTLVIMELMTIRGDLKSYLR SIRPEMENPYLAPPSISKINGMAGEIADGM AYLNANKFYHRDLAARNCMVAEDFTVKIGG GISNEOVLRFYWBGGLLDKPNVEPPHMEFILL MRMCWQYNPKMPSFLEIISSIKEEMEPGIRE VSFYYSEENKLPEPEDLLDFRWSYPLDPS ASSSSLP.PDRHSGHKAENGPPGVULVRASF DERQPYAHMNGGRKNERALPLPQSSTC USPRETISHE GRVFFRGGGVGHKGQULMGTFILAQDWLSE SNIFYCVSSMLRLOKRIASSVLRCGKKKW LDPNETNELAMNSRQQIRKGTAN ARMPEKYTWMSFRISKITARK ESKKIDRIMYHSLYLKVKGNYFKNKRILMEH HIELKADKARKKILADQARSKITEARK RRERLQAKKEEIIKTLSKEEETK SGAKASSSSRFTDSKSTKEARK RRERLQAKKEEIIKTLSKEEETK SGAKASSSSRFTDSKSTKEARK RRERLQAKKEEIIKTLSKEEETK SGDICNACVLULKRWKKLPAGSKKINWNH VVDARAGPSLKTTLKRYKKTLPAGSKKINWNH VVDARAGPSLKTTLKRYKKTLSGGRIHEVTS VSPRAGSSTSTSSGFGSG SGKKANAAEKEEGQGPEPLISTIGGCCT SGAKASSSSRFTDSKSTSGAKSSSRFTTSNSSSFNAGSLFTSNSSSFNA SSSSAGSSSVPYSSSSASAGSSPPTSNSSSSTNPAA SSSSAGSSSPYSSSSSAGSSPPTSNSSSSTNPAA SSSSAGSSSPYSSSSSSAGSSPPTSNSSSSTNPAA SSSSAGSSSPPTSNSSSSTNPAA SSSSAGSSSPYSSSSSAGSSPTSTNSSSSFTSSSGFSSSSSSSSSSSSSAGSSSPYSSSSSSAGSSPTSTNSSSSFTSSSFSSAGSSSSSSSSSSSSSSSSSSSSSSSSSS	i	}	1	1	ł		1
WDHRN.LTIK.AGKMYFAFNYKLCVSEIYMEE VTGTKGRQSKGDDYRNNGERASCESDV.HF TSTTTSKNRIIITWIR YRPPDYPDLISTYYK EAPFKNVTEYDGQDA.GGSINWIMVDVDLPP MKDVEPGILLIGLKPWTQYAVYVKAVTLTM VENDHIRGAKSEIL YIRTNAVSBIPLUVISAS NSSGLIVKWHYPSI.PINGLSYYIVZWQRQP QDGYLYRHNYCSKDKIPIKKYADGTIDIEVT ENPKTEVCQGEKGPCCACPKITAEKQAEKEE AFYRKVYENFI.HNSITYPRPERKRRDVMQVA NTITMSSRSRNTADTYNTIDELETTEYPFF ESRVDNKERTVISNLRPFILYRIDHISCNIHAE KLGCSASINEYARTMAPAGDIPGPYTWEP ESRVDNKERTVISNLRPFILYRIDHISCNIHAE KLGCSASINEYARTMAPAGDIPGPYTWEP PEPRISIL KWPEPENPINGLI MYEIKYGSQVE DQRECVSRQEYRKYGGAKINRLNPGNYTARI QATSLSGNGSWTIDPVFFYVQAKRYENFHILII ALPVAVLIVGGLVINLYVFHRKRINSRLGN GVLYASVINPEYSAADVYVPEWVAREKITI MSRELQQGSFGMYYEOVAKGVVKDEPTEV AIKTVNEAASMAREREFILARSVMKEPNCHH VVRILGVVSQGQPTI.VIMELMTRGDILKSYLR SURPPMENPIYLAPSLSKMIQMAGGIADGM AYLNANKFVRIDLAARNCMYAEDFTVKIGD FGMTRIDVETDYVARGGCKLDKYNWMSPESL KDGYTTYSDVWSPGVVLWEIATLAEQPYQ GLSNEQVLRFVNBGGGLLDKYPLOMENSEL KDGYTTYSDVWSPGVVLWEIATLAEQPYQ GLSNEQVLRFVNBGGGLLDKYPLOMENSEL KDGYPTYSDVWSPGVVLWEIATLAEQPYQ GLSNEQVLRFVNBGGGFGVVLVLRASF VSYYSEENKLPEPEELDLEPENNESSYLDDS ASSSSLPLPDRISGIKEANGEPGFGVVLVLRASF SASSSSLPLPDRISGIKEANGEPGFGVVLVLRASF ESSTYPCVSSMLRLQKRLASSVLRCGKKVW LPPRETNELANANSQQIRKKOTAN ARMPEKYTVWRRMGILRIPLAQDWLSE SNIFYCVSSMLRLQKRLASSVLRCGKKVW LPPRETNELANANSQQIRKKOTAN ARMPEKYTVWRRMGILRIPLAQDWLSE SNIFYCVSSMLRLQKRLASSVLRCGKKVW LPPRETNELANANSQQIRKKOTAN ARMPEKYTVWRRMGILRIPLAQDWLSE SNIFYCVSSMLRLQKRLASSVLRCGKKVW LPPRETNELANANSQQIRKKOTAN ARMPEKYTVWRRMGILRIPLAGPSKXVKM VVARAGPSLKTLIKKKKLPGSKKVMNMI VVARAGPSLKTTLKRKKKLPGSKKVMNMI VVARAGPSLKTTLKRKKVKLLPAGSKKVMNMI VVARAGPSLKTTLKRKKVKLLSGGNIKHSTICSCATI GGKASSSSSIRTPSKSKTERAR REBELQAKKRULLADQAEARRSKTKRAR REBELQAKKRULLADQAEARRSKTRAR REBELQAKKRULLADQAEARRSKTRAR REBELQAKCOTTLYRYFMSSMLYMI VVARAGPSLKTTLKRKVKTLSGGNIKHSTINSSTNPAA SSSASSSSVPOSSSSASSSSVPOSSSSASSSSVPOSSSSASSSSVPOSSSSASSSSVPOSSSSASSSSVPOSSSSASSSSVPOSSSSASSSSVPOSSSSSASSSSVPOSSSSASSSSVPOSSSSASSSSVPOSSSSASSSSVPOSSSSSASSSVPOSSSSSASSSVPOSSSSSS		1					
VITGKGRQSKGDDTRNGGRASCESDVLHF TSTITSKNRIITUHRYRPPDYBLISITVYVK EAPFKNVTEYDGQDACGSNSWAMVDVDLPP NKDVEPGILLBIGLRPWTQYAVYVKAVTLIM VENDHIRGAKSESLI YRITNASVSPIPLLVISAS NSSSGLIVKWNPPSLPNGRLSYYIVRWQRQP QDGYLYRHNYCSKDKEPRKYAPGTIDIEEVT ENKTEVCGGEKGPCCACPKTEAEKQAKEE AFYRKYFENI-HNSIPVPPBFRKRRDVMQVA NTIMSSRSRNTAADTYNITDPELLETEYPF ESRVDNKERTUSHNLRPFLYRIDHISCHTHAE KLGCSASNFVFARTMPAEGADDIPGPVTWEP RPFNSIFLKWPEPENNGLLIMYERKYGSQVE DQRECVSRQEYKKYGGAKINKINPGNYTAKI QATSLSGNGSWTDPVFTYVQAKKYEWHILLI ALPVAVLLIVGGUVINI-VYPHKRKNNSRLGN GVLYASVNPEYSAADVYVPDEWEVAREKIT MSRELQGGSFGMYTEGVAKGVVKDEPETRV AKTAVBEAASMERETELNEASVMKEPKCHH VVRLLGVVSGQPTL-IVMEIRGDLKSVLR SLRPEMENNPVLAPPSLSKMIQMAGELADGM AYLNANKFYHRDLAARNCAV/AEDFTVKIGD FGMTRIDVETDYYKRGKGLKPRWMSPESL KDGVFTTYSDVWSFGOVVLWEIATLACPYQ GLSNEQVLRFYWEGGLLKDKPNOPDMLFEL MRMCWQYNPKMEPSTLEISSIKEEMFGFRE VSSYYSEENKLPEPELDLEPENMESVPLDPS ASSSSLP.PDRHSGHKAENGPOPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC GRIFTGCGVGHKGQVLMGTFLAQDWLSE SNHYFCVSSMLRLOKRLASVLROGKKFW LDPNETNELAMNSRQQRKLENGLIRKPVT VHSRARCRNTLARRKGRHMGIGKRKGTAN ARMPEKYTWMFRMILIRLIKDGLIRKPVT VHSRARCRNTLARRKGRHMGIGKRKGTAN ARMPEKYTWMFRMILIRLIKDGLIRKPVT VHSRARCRNTLARRKGRHMGIGKRKGTAN ARMPEKYTWMFRMILIRLIKDGLIRKPVT VHSRARCRNTLARRKGRHMGIGKRKGTAN ARMPEKYTWMFRMILIRLIKDGLIRKPVT VHSRARCRNTLARRKGRHMGIGKRKGTAN ARMPEKYTWMFRMILIRLIKDGLIRKPVT VHSRARCRNTLARRKGRHMGIGKRKGTAN ARMPEKYTWMFRMILIRLIKDGLIRKPVT VHSRARCRNTLARRKGRHMGIGKRKGTAN ARMPEKYTWMFRMILIRLIKDGLIRKPVT VHSRARCRNTLARRKGRHMGIGKRKKTWM VVDRAGPSLLKTUKKKKLPAGSKKNWMH VVDRAGPSLLKTUKKKKLLADAGSKKNWMH VVDRAGPSLKTTLKKRKKTLAGSKKNWMH VVDRAGPSLKTTLKKRKKTLAGSKKNWMH VVDRAGPSLKTTLKRKKKTLAGSKKNWMH VVDRAGPSLKTTLKRKKKTLAGSKRKNWMH VVDRAGPSLKTTLKRKKKTLAGGRANDRUFF TVNORFWTGSDTTOSGTFGCSCCHETTA GGDICANACVLLUKRWKLLAGSKKNWMH VVDRAGPSLKTTLKTRKKKTLAGGRANDRUFF TVNORFWTGSDTTOSGTFGCSCCHETTA GDICANACVLLUKRWKLLAGSKKNWMH VVDRAGPSLKTTLKTRKWTTLSGGRCCC SNKKANAAEKEEGQGPEPLISTGCWTTEVFM TVMCRGKCCGONYKGRGGVLDTHLFKPCC SNKKANAAEKEEGGGPLDTHTAAAGGGATPG NAGPLAPNGSAAAPFAGSARNFTSNSSSTNPAA SSSAGSSSVPYSSSSAAGSVPYSSSSSASSSSTNPAA SSSAGSSVPYSSSSAAGSVPYSSSSASSSSSTNP		ļ				1	1
TSTITSKNRIIITWHRYRPPDYRDLISFTVYYK EAPPKNNTEVDGQDACSWNMMVDVDLPP NKDVEPGILHGLKPWTQYAVYVKAVTLTM VENDHIRGAKSEL YIRTNASVPBIPLOVLSAS NSSSQLIVKWNPSLPNGNLSYYTVRWQRQP QDGYLYRHNYCSKDKIPIRKYADGTIDEEVT ENPKTEVCGGEKGPCACKFEAERQAEKEE AFYRKVFENFLHNSIFVPRPERKRRDYMQVA NTIMSSRSRNTTAADTYNTIDPEELETEYPF ESRVDNKERTVISNLRPFILYRIDHSCNHEAE KLGCSASNFVFARTMPAGADDIPGYTWEP RPENSIFLKWFEPENPYRGLILMYEIKYGSQVE DQRECVSRQEYRRYGGAKLNRLNPGNYTARI QATSLSGNGSWIDTPVFYDEWEVAREKIT MSRELQQSFGMVYEGVAKGVVKDEPETRV AIKTVNEAASMREIFLNEASVMKEPNCHH VVRLLGVVSQQPTLVMBLMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGELADGM AYLNANKFVHRDLAARNGMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVPTTYSDVWSFGVLVLARSF EVFYYSEENKLPEPELDLEPENMESVPLDPS ASSSLPLPDRHSGHKAENGPGPCVLVLRASF EVFYYSEENKLPEPELDLEPENMESVPLDPS ASSSLPLPDRHSGHKAENGPGPCVLVLRASF DERQPYAHMNGGRKNEAFLPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGVLMGTFILAODWLSE SNHYFCVSSMLRLOKRLADDWLSCKKWW LDPNETNEIANANSRQJIRKLIKDGLIIRRPVT VHSRARCRKNTLARRKGRHMGIGKRGTAN ARMPEKVTWMRRMRILRRLRRYRESKRYR ESKIDRHMYHSLYLKVKGNVFKNRLIMEH HIKLKADKARKLLADDQAARRSKTKEARK REEERLQAKKEEIKILSKEEFKK SGDICNACVLLLLKRWKKLPAGSKKNWHN VVDARAGFSVLKTILKKKVETLSSONRIKIST QISKLQKEFKRHNSDAHSTITSISASPAQSFLF TVNGFRWTGSDTGGTFGSNRNIPVFSFLDLA TYWKRQKICCGIVYKGRFGEVLIDTHLFKPCC SNKKANAAEKPEEQGFEDISTOEWVTEVPCM SSSAGSSSVPVSSSAASARGIGGISTTSSSGFGSS	1	İ	i	! .			l
EAPFKNYTEYDGQDAGGSNSWNMDVDULPP NKDVEPGILLHGI.KPVTQAVYVKAVTLTM VENDHIRGAKSEILYIRTNASVPSIPLDVLSAS NSSQLJVKWNPPSLPNGNLSYTYNWQRQP QDGYLYRHIYNCSKDKIPFIKYADGTIDIEEVT ENPKTEVCGGEKGPCCACPKTEAERQAEKEE AEYRKVFENFILHNSIFYPFERKRDDWQVA NTTIMSSRSRNTTAADTYNITIDPEELETEYPFF ESRVDNKERTVISNLRPFILYRDIHSCNHEAE KLGCSASNFVFARTIMPAGADDIPGPVTWEP RPENSIFLK.WPEPENPNGLLLMYEIKYGSQVE DQRECVSRGYPKRYGGAKLNRLNIPGNYTTARI QATSLSGNGSWTDPVFFYVQAKRYENFIHLII ALPVAVLLIVGGLVMLIVIGGLVMLVYFIKRKNNSRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELQGSFGMVYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFILNEASVMKEPNCHH VVRILIGVSQQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMOMAGEIADGM AYINANKFVIRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVYWEIATLAEAPPYQ GLSNEQVLRFVMEGGLLDKFDNCPDMLFEL MRCWQVNPKMRPSFLEISSIKEEMEPFGRE VSFYYSEENKLPEPELDLEPEMBESVPLDPS ASSSSLPLDERHSGHKARMQMAGEIADGM AYINANKFVIRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGFGCVUMGTEILAODWLSE SSSSLPLDERHSGHKARHQMAGGIADGM AYINANKFVIRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGFGCVUMGTEILAODWLSE SSSSLPLDERHSGHKARHQMAGGIADGM AYINANKFVIRDLAARNCMVAEDFTVKIG BERDEN SSSSLPLDERHSGHKARHQMAGGEINGFVVT VFISHARCKKVTLLAGDWLSE SSSSLPLDERHGHKARHGIGKRKGTAN ARMPEKVTWMRRRRILRELLREYMESKRYR ESKKIDRHMYHSLYLKVGNOVFKNKRILMEH HIELKADKARKKLLADQAEARRSKTKEARK REERLQAKKEEIKTLSKEETEKK SGDICNACVLLILKRWKKLPAGSKKINWNH VVDARAGPSUKTTLKFKKVTLISGNRIKIST QISKLQKEFKRHNSDAHSTTISSASPAQSPLF TVNOFRWTGSDTGGTPGSNRNIPVSFGLDL TYWKRQKICCGIUYKGRFGEVLIDTLFFPCC SIKKANAAREKPEEQGPEPLISTIGWTEVFM 823 2173 A 6727 3 4663 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAAPAGSAARNFTSNSSTNPAA SSSSASSSSPVFOSSASAARNFTSNSSTNPAA			ļ				,
NKDVEPGILHGI, EWPTQYAVYVKAVTLTM VENDHIRGAK SEIL VIRTNAS YPSPIPLOVI, SAS NSSQLIVK WNPPSLPNGNLSYYIVR WQRQP QDGYL YRHNYCSKOKIPIRK YA QGTDIEEVT ENPKTEVCGEKGPCCACKTEA EKQAEKEE AFYKK VFENFLHNSIFVPREPRK RENDYMQVA NTTMSSRSRNITTAAD TYNTIDPEELETE YPFF ESRYDNKERTVISNIR PFTLYRDIHSCNHEAE KLGCS ASNFVP ARTIMPAGDD TEP YWFP RPENSIFLK WPEPENPNGLILMYEIK YGSQVE DQRECVSRQEYRK YGGAKLNKLINFGNTTABI QATSLSGNGSWIDP VFTYDE WEVAREKIT MSRELQGOSFGMVYEGVAK YEVENFHILLI ALPVAVLLIVGGL VIMLYVFHRKRNISKLGN GVLVAS VINPEYFSAAD YVAK YEPENYNSKLGN WSTALGOSFGMVYEGVAK GVK KUPK PREVENTHILLI ALPVAVLLIVGGL VIMLYVFHRKRNISKLGN GVLVAS VINPEYFSAAD YM KEPNCHH VVRLLGVVSQOPTL VIMBLIMTRGDLK SYLR SLRPEMENNFVLAPPSLSKM QMAGELADGM AYLNANK FVHEDLAAR ANCHVAEDFTVKIGD FGMTRDIYETDYYRK GGKGLLFVR WMSPESL KDGVFTTYSDVWSFGVLVLARASPETLEISSKE EMERPGFRE VSFYYSEENKLPEPELDLEPENMES VPLDPS ASSSLPLPDRHSGHKAENGAPPOLVLARASF EVSFYYSEENKLPEPELDLEPENMES VPLDPS ASSSLPLPDRHSGHKAENGAPPOLVLARASF EVSFYYSEENKLPEPELDLEPENMES VPLDPS ASSSLPLPDRHSGHKAENGAPPOLVLARASF EVSFYYSEENKLPEPELDLEPENMES VPLDPS ASSSLPLPDRHSGHKAENGAPPOLVLARASF EVSFYYSEENKLPEPELDLEPENMES VPLDPS ASSSLPLPDRHSGHKAENGAPPOLVLARASF EVSFYYSEENKLPEPELDLEPENMES VPLDPS ASSSLPLPDRHSGHKAENGAPPOLVLARASF EVSFYYSENKLASSVLRCGKKKW LDPNETNEIANANSR QQURKLIKDGLIRRPYT VHSRARCKNTLLARRKGRHMGIGKRKGTAN ARMPEK VTWMRRMRILRLRRYRESKRYR ESKLIDRHMYHSLYLKVKGNVFKKRILMEH HEILKADKARKLLADQAGARRSKTKEARK REEFILQAKKEEIGHFFYKWRIPSEGCCI SGAKSSSSRFTDSKRYEKUPGOSKRINWHY VVDARAGFFHEYKKMFFFFHEYKMYRSIEGCCI SGAKSSSSRFTDSKRYEKUPGOSKRINWHY VVDARAGFSULKTILKRKVKTLISGNRIKIST QISKLQKEFKRIHNSDAHSTTISSASRAQSFLF TVNOFRWTGSDTGGUFFGSNRINFVSFILDLA TYWKRQKICCGIUYKGRFGEVLIDTHLFFPCC SNKKANAAEKPEEQGFEPLISTOEWVTEVFM SSSAGSSSVPVSSSAASARDISGISTTSSSSFIDSA SSSAGSSSVPVSSSAASARDISGISTTSSSSFIDSA SSSAGSSSVPVSSSSAASROIGGSTTSSSSGFGS			1	1			
VENDHIRGAKSEL YURTMASVPSIPLOVLSAS NSSQLIVE WAPPSLPHORLSYTVTWOROP QDGYLYRHNYCSKDKIPIKYADGTIDIEEVT ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE AEYRKVPENFLHINSTPHYPERKRRDYMQVA NITMSSRSNITIAADTYNITDPEELETEYPFF ESRYDNKERTVISNLEPTLYRDIHSCNIEAE KLGCSASNFVFARTMPAEGADDIGPVTWEP RPENSIFLKWPEPENPNGLLIMYGENTYARI QATSLSGNGSWTDPVFFYVQAKRYENFIHLI ALPVAVLUVGGLVMLIVYFHIRKRNNSRLGN GVLYASVNJEYSPAADVYYPDEWEVAREKTI MSRELGQGSFGMYEVGVKGVVKDEPETRV AIKTVNEAASMRERIEFLNEASVMKEPNCHH VVRLIGVVSQGPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLAENOCHMENTERGLKSYLR SLRPEMENNPVLAPPSLAENOCHMENTERGLKSYLR SLRPEMENNPVLAPPSLEIDSKEEMEPGER VSFYYSEENKLPEPELDLEPENMESVPLDPS ASSSLPLPPRHSGHKANGPGGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTHAQDWLSE SNHYFCVSSMLRLOKRLASSVLRCGKKKVW LDPNETNEIANANSRQQRKLIKGGLIBVRYMS ESKLIDRHMYHSLYLKVKGNVFKNKRILMEH HIKLKADKARKKILADQAEARRSKTKEARK REERLQAKKEEIKTLSKEEETKK BERELQAKKEEIKTLSKEETKK BERELQAKKEEIKTLSKEETKK QISKLQKEFKRHNSDAHSTTSASSPAQSPLC TVNQFRWTOSDTGVGFGGSNRNHPVSSEDLIL TYWKRQKICCGRIYKGRFGGVLDTHLFKPCC SIKKAVAAEKPEEQGPPEPISTOGEWTEVTEVTM VVDARAGPSLKTTLKPKKVKTLSGNRIKIST QISKLQKEFKRHNSDAHSTTSASSPAQSSLDLT TYWKRQKICCGRIYKGRFGGVLDTHLFKPCC SIKKAVAAEKPEEQGPPEPISTOGEWTEVTEVTM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATFG NAGPLAPNGSAAPPAGSAFNISTISNSSSTNPAA SSSASGSSVPYNSSSASGSIGSITSTSSSSTSSGFGSS		ļ	1	ł	}		EAPFKNVTEYDGQDACGSNSWNMVDVDLPP
NSSSQLIVKWNPPSLPNGNLSYYTVRWQRQP QDGYLYRHNYCSKDKIPIRKYADGTIDIEEVT ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE AEYRKVFENFLHNSIFVPRPEERRRDVWQAA NTIMSSRSNITAADTYNITDFELETEYPFF ESRVDNKERTVISNLRPFTLYRIDHSCNIEAE KLGCASNFVFARTMPAEGADDIPGPVTWEP RPENSIFLKWPEPENFINGLILMYEIKYGSQVE DQRECVSRQEYRKYGGAKLINRLNPGNYTARI QATSLSGNGSWTDPYVQARKYENFHILII ALPVAVLLIVGGLVIMLYVHRKRNNSRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMVYEVVAKEVKDEPETRV AIKTVNEAASMRERETELNEASVMKEFNCHH VYRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGELADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVLWEISTLAEQPYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSLPLPDRHISGHKAENGPGPGVLVLRASF DERQPYAHMNGGKNERALPLPQSSTC GRVLFRGGGVGHKGVUMGTFILAQDWLSE SNHVFCVSSMLRLOKRLASSVLRCGKKKVW LDPNETNEIANANSRQURKLIKDGLIRKPVT VHSRARCRKNTLARRKGRIMGIGKRGTAN ARMPEKVTWMRRMRILRRLRRYRES/KRYK ESKUDDRHMYHSLYLKKONYFKNKRILMEH HKLKADKARKKLLADQABARSKITKEARK REERLQAKKEEIKTLSKEETTKK 822 2172 Å 6715 772 21 DFRFGLLLPRKKKMFHKMYRESCCCI SGAKSSSSRFTDSKRYEKUPGSCFGLHETR SGDINCNACVLLLIKRKVLPAGSKKNWMH VVDARAGPSUKTTLKPKKVYTLSGNRIKSTF TVNGFRWTGSDTGVGFPGSNRNHPVSSEDLL TYWKRQKICCGRIYKGRFGEVLIDTHLFKPCC SIKKAVAAKEREEGGPFPLPISTOEWVTEVFM 823 2173 Å 6727 3 4663 PYLATLQLDSSLLIPPKYQTPPAAAQGATFG NAGPLAPNGSAAPFAGSAFNISTNSSSTNPAS SSSAGSSSVPYNSSSAAPGISGUSTTSSSGFGS	1	1	l				NKDVEPGILLHGLKPWTQYAVYVKAVTLTM
NSSSQLIVKWNPPSLPNGNLSYYTVRWQRQP QDGYLYRHNYCSKDKIPIRKYADGTIDIEEVT ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE AEYRKVFENFLHNSIFVPRPEERRRDVWQAA NTIMSSRSNITAADTYNITDFELETEYPFF ESRVDNKERTVISNLRPFTLYRIDHSCNIEAE KLGCASNFVFARTMPAEGADDIPGPVTWEP RPENSIFLKWPEPENFINGLILMYEIKYGSQVE DQRECVSRQEYRKYGGAKLINRLNPGNYTARI QATSLSGNGSWTDPYVQARKYENFHILII ALPVAVLLIVGGLVIMLYVHRKRNNSRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMVYEVVAKEVKDEPETRV AIKTVNEAASMRERETELNEASVMKEFNCHH VYRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGELADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVLWEISTLAEQPYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSLPLPDRHISGHKAENGPGPGVLVLRASF DERQPYAHMNGGKNERALPLPQSSTC GRVLFRGGGVGHKGVUMGTFILAQDWLSE SNHVFCVSSMLRLOKRLASSVLRCGKKKVW LDPNETNEIANANSRQURKLIKDGLIRKPVT VHSRARCRKNTLARRKGRIMGIGKRGTAN ARMPEKVTWMRRMRILRRLRRYRES/KRYK ESKUDDRHMYHSLYLKKONYFKNKRILMEH HKLKADKARKKLLADQABARSKITKEARK REERLQAKKEEIKTLSKEETTKK 822 2172 Å 6715 772 21 DFRFGLLLPRKKKMFHKMYRESCCCI SGAKSSSSRFTDSKRYEKUPGSCFGLHETR SGDINCNACVLLLIKRKVLPAGSKKNWMH VVDARAGPSUKTTLKPKKVYTLSGNRIKSTF TVNGFRWTGSDTGVGFPGSNRNHPVSSEDLL TYWKRQKICCGRIYKGRFGEVLIDTHLFKPCC SIKKAVAAKEREEGGPFPLPISTOEWVTEVFM 823 2173 Å 6727 3 4663 PYLATLQLDSSLLIPPKYQTPPAAAQGATFG NAGPLAPNGSAAPFAGSAFNISTNSSSTNPAS SSSAGSSSVPYNSSSAAPGISGUSTTSSSGFGS			1		i		VENDHIRGAKSEILYIRTNASVPSIPLDVLSAS
ODGYĹYRIHYCSKOKUPIRKYADGTIDIĒCY ENPKTEVGGĒKEROCAPKTEARKQAĒKĒ AEYRKVFENĪLINSIFVPRPĒRKRDVMQVA NITIMSSRSRNĪTAADTYNĪTIPĒLĒTĒVPPF ESRVDNKĒRTVISNLIRPĪLIYRDĪHSCNĪĒAE KLGCSASNFVFARTMPABGADDĪPGPVTWĒP PĒPSIFĪL WPPĒPSNPĪGLĪMYĒLKĪVGSQVĒ DQRĒCVSRQĒYRKYGGĀKLNRLNPGNYTĀRĪ ALPVĀVLLĪVGGLVIMĪLYVĒHKSKNISRLGĪ GVLYĀSVNPĒVJFGVĀKĀKYBENĪĒLĪL ALPVĀVLLĪVGGLVIMĪLYVĒHRKENNSRLGĪN GVLYĀSVNPĒVJFGVĀKĀKYBENĪĒLĪL MĀRĒLĢGSFGMYĀDĀVĀVPDĒWĒVĀREŅĪT MSRĒLĢGSFGMYĀDĀVĀKYDĒPĒTRV ĀIKTVNĒĀASMĀRĒRĪĒHNĒASVMKĒPĪCHĒH VVRLLGVVSQQPTLVĪMEMMTROLKSYLR SLRFĒMĒNNPVLĀPPSLSKMĪQMĀGĒLĀDGM ĀYLNĀNKFVHRDLĀĀRNCMVĀEDFĪVKĪGG FGMTRDĪVĒTDYYRKĢGLĪPĀVRWĀPĒSLE KDGVFTTYSDVWSFGVVLWĒJATLĀĢDPVQ GLSNĒQVLRFVMĒGGLĪJĀPDNCPDMLĪĒL MĒMCWQYNPKMĒPSĪLEĪSSIKĒEMĒPĢFRE VSFYYSĒNKLPĒPĒLDLĒPĒNMĒSVPLDPS ĀSSSLPĪPĀRISGĪKAROPĢFGVLVLRĀSF DĒRQPYĀHMNĪGGRKNĒRĀJĒLPQSSTC SNHYPCVSSMLRLQĶRLĀSSVLRCĢKKKVW LDPNĒTNĒLĀNANSRQQĪKKLĪKGOLĪJRKPVT VHSRĀRCRNĪTLĀRRĀRIMGĪGKRKGTĀN ĀRMĒKŪDRIMĀHSTLĀKRYRĒSKRYR ĒSKLĪDBRIMYHSLYLKĢNVŠKNRĪLMĒH HKLKADKARKKLADQĀĒARSKTKĒĀRK RĒĒRLQĀKKĒLĪKLĀKETĪKK 822 2172 Ā 6715 772 21 DFRFGLLIPĀRKKKVKONVĒKNIKLMĒH HKLKADKARKKLLADQĀEARRSKTKĒĀRK RĒĒRLQĀKKĒLĪKIKĢRĒKĶRMYRĪSĒĞCCCĪ SGAKSSSSRFĪDSKRYĒKKŪMĢGRKGCĪĞLĪTH VVDARAGPSUKTĪLKPKKVKTLSGNRĪKUST QISKLQKĒFKRHNISDĀHSTTSASSĀROPSIEDLĀ TYWRQKICCGŪNYKĢRĢĒGVLIDTHLFKPCC SNKKĀJĀĀKĀRĀPĒĒGGORPĢGSNRNHPVFSĒLDLĀ TYWRRQKICCGŪNYKĢRĢĒGVLIDTHLFKPCC SNKKĀJĀĀKĀRĀPĒĒGGORĪGOSISTISSSSTĪPĀA SSSASGSSVPPVSSSASARĪGSISĪSTISSSSTSPA	}		1				
ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE AEYRKVFENFLHNSIFVPRPERKRRDYMQVA NTIMSSRSNTTAADTYNITDEELETEYPFF ESRVDNKERTVISNLRPFILTRIDIHSCNHEAE KLGCSASNFVFARTMPAEGADDIPGPYTWEP RPENSIFLKWPEPENPNGLILMYEIKYGSQVE DQRECVSRQEYRKYGGAKLNRLNPGNYTARI QATSLSGNGSWIDVFFYVQAKRYENFHLII ALPVAVLLIVGGL VIMLYVHRKRRNSRLGN GVLYASVNPEYFSAADVYVPDEWVAREKITI MSRELGQGSFGMVVEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSQGQPILVIMELMTRGDLKSYLR SLRPEMENNFVLAPPSLSKMIQMAGELADDM AYLNANKFVIRBLDAARNCMVAEDFTVKIGD FGMTRDIYSTDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVVMEGGLLDKPDNCPDMLFEL MRMCWQYNFKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLFEDLEFENMESVPLDPS ASSSSLPLPDRHSGHKAENGPOFGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHRGGVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIRKPVT VHSRARCRKNTLARKGRHMGIGKRKGTAN ARMPERVTWMRRMRILRRYRESKRYR ESKLIDAKSVLRCGKKKVW UDPNETNEIANANSRQQIRKLIKDGLIRKPVT VHSRARCRKNTLARKGRHMGIGKRKGTAN ARMPERVTWMRRMRILRRYRESKRYR ESKLIDAKAKLLADQAARRSKITKEARK REBELQAKKEEIKITJSKEEETKK 822 2172 A 6715 772 21 DFRFGLLPFRKKMFGFHKFKMYRSIEGCCI SGAKSSSISTTDSKRYEKUPQSCFGLHETRI SGDNCANCVLLUKRWKKLPAGSKKNWMI VVDARAGPSUKTTILKPKEVKTLJSGNRIKST QISKLQKEFKRHNSDAHSTTSNSASPAQSPLF TVNQFRWTGSDTGVGFGSNRNHPVFSSFLDLI TTVNGFRWTGSDTGVGFGSNRNHPVFSSFLDLI TVNGFRWTGSDTGVGFGSNRNHPVFSSFLDLI TVNGFRWTGSDTGVGFGSNRNHPVFSSFLDLI TVNGFRWTGSDTGVGFGSNRNHPVFSSFLDLI TVNGFRWTGSDTGVGFGSNRNHPVFSSFLDLI TVNGFRWTGSDTGVGFGSNRNHPVFSSFLDLI TVNGFRWTGSDTGVGFGSNRNHPVFSSFLDLI TVNGFRWTGSDTGVGFGSNRNHPVFSSFLDLI TVNGFRWTGSDTGVGFGSNRNHPVFSSFLDLI TVNGFRWTGSDTGVGFGFPLISTTGWTTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPFKYQTPPAAAQQATFG NAGPLAPNGSAAPPAGSAAPPGASARTPTSNSSSTNPAA SSSASSSSVPVSSSASAPGISQISTTSSSSFSSS	1						
AEYRKVFENPLHNSIFVPRPERKRRDÝMQVA NTTMSSRSRNITAADTYNITDPEELETEYPFF ESRVDNKERTVISNLRPFTLYRIDHISCNHEAE KLGCSASNFVFARTMPAEGADDIPGPVTWEP RPENSIFLKWPPEPENPNGILMYEIKYGSQVE DQRECVSRQEYRKYGGAKLNRLNPGNYTARI QATSLSGNGSWTDPVFFYVQAKRYENFIHLII ALPVAVLLIVGGLVIMLYVPHRKRNNSRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMVYEGVAKGVVKDEPETRY AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSGQPTLVIMEMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIGMAGEIADGM AYLNANKFVIRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGKGKGLLPVRWMSPESL KDGVVFTTYSDVWSFGVVLWEIATLAEQFYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRFSFLEISSIKEEMPEGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGFQVLVLRASF DERQPYAHMNGGREADGM VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGFQVLVLRASF DERQPYAHMNGGREADFA VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGFQVLVLRASF DERQPYAHMNGGREADFA VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKERALPLPQSSTC SQRVLRGCGVGHKGQVLMGFFLAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRGTAN ARNPEK VTWMRRMLIRLLRRYRESKRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HIKLKADKARKKLLADQAEARRSKTKEARK REPELQAKKEEHIKLISKEEPETK SQDINCMA(CVLLLKRWKKLPAGSKKNWMH VVDARAGPSULKTTLKPKKWTLISGNRIKST QISKLGKEFKRWHNSDAHSTTSISASPAGSFLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC SNKKAMAEKPEGGPEPLPISTQEWYTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTTNPAA SSSASGSSVPPVSSSASAPGISQISTITSSSGFSGS		ļ	1				l '
NTTMSSRSRNTTAADTYNITDPEELETEYPF ESRVDNKERTVISNLERPTILYRIDHSCNHEAE KLGCSASNFVFARTMPAEGADDIPGPVTWEP RPENSIELK WPEPENPNGLILMYEIKYGSQVE DQRECVSRQEYRKYGGAKLNRLINPGHYTAKI QATSLSGNGSWTDPVFFYVQAKRYENFIHLII ALPVAVLLIVGGL VIMLYVYHRKRRINSRLGN GVLYASVNPEYFSAADVYVPDEWVAREKITI MSRELQQGSFGMVYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFLINEASVMKEFNCHH VVRLLGVVSQGQPTLVIMEIMTRGDLKSYLK SLRPEMENNPVLAPPSLSKMIQMAGELADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIVETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEGVLRFVVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSLPLPDRHSGIKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGGGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLLRFYRESÆKYR ESKKIDRIMYHSLYLKVGGNYFKNRILMEH HKLKADKARKKLLADQAEARRSKTKEARK REERLQAKKEEIIKTLSKEETTK 822 2172 A 6715 772 21 DPEPGLLLPRKKKMFGFHKFKMYRSIEGCCI SGAKSSSSFTSVENSKSFGCLHETRI SGDICNA\CVLL\LKRWKKLPAGSKK\NWH VVDARAGPSUKTTLKPKKVTLSIGNRIKST QISKLQKEFKRINSDAHSTTSSASPAQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDLA TYWKRQKICCGNIYKGRFGCVLIDTHLFKPCC SNKKANAAEKPEGOFEPLPISTGEWYTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSSASAPGISQISTITSSSGFSGS			į]			
ESRVDNKERTVISNLRPFTLYRIDHISCNHEAE KLGCSASNFVFARTMPAEGADDIPGPVTWEP RPENSIFLKWPEPENNGLLIMYEIKYGSQVE DQRECVSRQEYRKYGGAKLNRLNPGNYTARI QATSLSGNGSWIPVFFYVQAKRYENFHLII ALPVAVLLIVGGLVIMLYVFHRKRNNSRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMYYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMQMAGELADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGG FGMTRDIYETDYRGGKGLIDKPDNCPDMLFEL MRMCWQYNPKYRGGKGLIDKPDNCPDMLFEL MRMCWQYNPKYRGGKGLIDKPDNCPDMLFEL MRMCWQYNPKYRGGKGLIDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHIMNGGKNERALPLPQSSTC SPIPYFCVSSMLKAENGPHMGIGKKFVT VHSRARCRKNTLARRGHMGIGKKGTAN ARMPEKVTWMRRMRILRFILRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HKLKADKARKKLLADQAEARRSKTEARK RREERLQAKKEEIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLIPRKKKMFGFHKPKMYRSIEGCCI SGAKSSSSVFTDSKRYEKUPGVSCFGLHETR; SGDICNAICVLLLKRWKKLPAGSKKNWH VVDARAGPSLKTTLKPKKVKTLSGNBIKIST QISKLQKEFKKHINSDAHSTTSSASPAQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDLA TYWKRQKICCGNYKGRFGEVLIDTHLFKPCC SIKKKAAAEKPEGGPBPLPISTQEWYTEVYEM 823 2173 A 6727 3 4063 PYLATLQLDSSLIPPKYQTPPAAAQGOATPG NAGPLAPNGSAAPPAGSAPPAGSAPPITSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSSFSGS	•	i					•
KLGCSASNFVFARTMPAEGADDIPGPVTWEP RPENSIFLK WPEPENPNGLILMYEIKYGSQVE DQRECVSRQEYKYGGAKLNRLNPGNYTARI QATSLSGNGSWTDPVFFYVQAKRYENFIHLII ALPVAVLLIVGGLVIMLYVFHRKRNNSRLGN GVLYASVNPEVFSAADVYVPDEWEVAREKIT MSRELGQGSFGMYYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSQGPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGELADGM AYLNANKFVHRDLAARNGMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRISGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHYFCVSSMLRLQKRLASSVLRCGKKKW LDPNETNEIANANSRQQIRKLLKDGLIRKPVT VHSRARCKKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVGNVFKNKRILMEH HIKLKADKARKKLLADQAEARSKTKEARK REERLQAKKEEIIKTLSKEEETKK SGDICNAICVLLLKRWKKLPAGSKKNWNH VVDARAGPSLKTTLKRKVKTLLSGNRIKST QISKLQKEFKKWINSDAHSTTSSASPAQSPLF TVNQFRWGGSDTGJFFFSNRIHPYFSFLDLA TYWFRQKICCGNIVKGRFFKKWTLSGNRIKST TVNQFRWTGSDTGVGFFGSNRIHPYFSFLDLA TYWKRQKICCGNIVKGRFFKKWTLSGNRIKST TVNQFRWTGSDTGVGFFGSNRIHPYFSFLDLA TYWKRQKICCGNIVKGRFFCEVLIDTHLFKPCC SNKKAMAAEKPEEQGPEPLPISTQEWVTEVFM RS23 2173 A 6727 3 4063 PYLATLQDSSLLPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAPPAGSASPNTSNSSTNPAA SSSASGSSVPPVSSSASAFGISQISTTSSSSFGSS	l	l	i	ł		ł	l
RPENSIFLKWPEPENPIGILIMYEIKYGSQVE DQRECVSRQEYRK YGGAKLNRLNPGNYTARI QATSLSGNIGSWTDPVFFYVQAKRYENFIHLII ALPVAVLLIVGGLVIMLYVFHKRNNSRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMYYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENRPVLAPPSLSKMIQMAGEIADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSSYYSEENKLPEPEELDLEPENMESVPLDPS ASSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFLAQDWLSE SNHYFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEK VTWMRRMILRRILRRYRES/KRYR ESKKIDRHMYHSLYILKVKGNVFKNKRILMEH HIKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEBIIKTLSKEETIKK 822 2172 A 6715 772 21 DFRPGLLLPRKKKNFGFHKPKMYRSIEGCCI SGAKSSSSLRTDSKRYENDFQSCCFGLHETRI SGDICNAICVLLUKRWKKLPAGSKKINWNHI VVDARAGPSLKTTLKPKKVKTLSGNRIKST QISKLQKEFKRHINSDAHSTTSSASPAQSPLF TVNQFRWTGSDTGVGFPGSNRNHEVFSFLDLA TYWKRQKICCGIVKYGRPEVLIDTHLFKPCC SNKKANAAEKPEDQGPEPLPISTQEWYTEVFM 823 2173 A 6727 3 4063 PYLATLGDSSLLPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFPITSNSSSTNPAA SSSASGSSVPYSSSASAFGISQISTTSSSGFSGS							
DQRECVSRQEYRKYGGAKLNRLNPGNYTARL QATSLSGNGSWTDPVFFYVQAKRYENFIHLII ALPVAVLLIVGGLVIMLYVFHRKRNISRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMVYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFINEASVMKEFNCHH VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGELADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC VSFYSEENKLPEPEELDLEPENMESVPLDPS ASSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKINERALPLPQSSTC VSFYSEENKLINGAGRKGVLMGHKKVV UDPNETINEIANANSRQQIRKLIKDGLIIRKPYT VHSRARCRKNLARKKILLAQDAEARRSKTKEATK ESKKIDRHMYHSLLKVKGNVFRKFKILMEH HIKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIKILSKEEETKK RREERLQAKKEEIKILSKEEETKK RREERLQAKKEEIKILSKEETKK SGDNCNACVLLLLKRWKKLPAGSKKINWNH VVDARAGPSULKTILKPKVKTLISGNRIKST QISKLOKEFFKRHINSDAHSTITSUSASPAQSPLF TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TYNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGVGFPGSNRINIPVFSFLDLI TVNQFRWTGSDTGGVGFPGSNRINIPVFSILDSSGFSGS SSSAGSSVPVSSSASAPGISQISTTSSSGFSGS		1			į	i	l
QÁTSLSGNGSWTDPVFFVVQAKRYENFIHLII ALPVAVLLIVGGLVIMLYVFHKKRNNSRLGN GVLYASVNPSYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMVYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGELADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTEDTYSTLYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVNMEGGLLDKFDNCPDMLFFL MRMCWQYNPKMRPSFLEIISIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERAHPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHYFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLINGGLIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMFEKVTWMRRMFLIRRLLRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIKKTJSKEEETKK BEZ 2172 A 6715 772 21 DFRFGLLLPRKKKMFGFHEKPMYRSIEGCCI SGAKSSSSRFTDSKRYEK\DFQUSCFGLHETRI SGDICNAICVLLLLKRWKKLPAGSKKINWNH VVDARAGPSKLTLIKKPKWTNSIEGCCI SGAKSSSSKFTDSKRYEK\DFQUSCFGLHETRI SGDICNAICVLLLLKRWKKLPAGSKKINWNH VVDARAGPSKLTLIKKVEKVTLISGRIKIST QISKLQKEFKRHINSDAHSTTSISASPAQSPLF TVNOFRWTGSDITGVGFGSSNRNIHPVFSFLDLI TYWKRQKICCGNIYKGRGFGEVLIDTHLFKPCC SNKKAIAAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLOLDSSLLPPRYQTPPAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSAGGSVPVSSSASAPGISQISTTSSSGFSGS	l	1	i	ł			1
ALPVAVLLIVGGLVIMLYVFIRKRNNSRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMYFEGVAKGVVKDEPETRY AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSQGGPTLVIMELMTRGDLKSYLR SLRFEMENPVLAPPSI SKMIQMAGEIADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHISGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHYFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEK VTWMRRMILRILRYRSJKRYK ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH IHKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRFGLLJPRKKKMFGFHKPKMYRSIEGCCCI SGAKSSSSRFTDSSRYEKUDFQUSCFGLHETR\ SGDICNA\CVLLULKRWKKLPAGSKK\NWNH VVDARAGPSLKTTLKPKKVKTLSGNRIK\ST QISKLQKEFKR\HNSDAHSITS\SASPAQSPLF TVNQFRWTGSDITGVGFFGSVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTFPAAAQGQATFQ NAGPLAPNGSAAPPAGSAPNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	İ		ļ				DQRECVSRQEYRKYGGAKLNRLNPGNYTARI
GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMVYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFILNEASVMKEFNCHH VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGEIADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGG FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVPTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVMBGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPDHISGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQURKLIKDGLIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLRRYRES/KRYR ESKKDRHMYHSLYLKVKGNVFKNKRILMEH HHLIKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEETTKK 822 2172 A 6715 772 21 DFRFGLLLPRKKKMFGFHKPKMYRSIEGCCCI SGAKSSSSRFTDSKRYEKUDFQISCFGLHETRI SGDIICNAICVLLUKRWKKLPAGSKKINWNH VVDARAGPSLKTTLKPKKVKTLISAGSRIKIST QISKLQKEFKRWINSDAHSTTSISASPAQSPLF TVNQFRWTGSDTGVGFFGSNINHPVFSFLDLI TYWKRQKICGGNYKGRFGEVLIDTHLFKPCC SNKKAIAAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTFPAAAQGQATFG NAGPLAPNGSAAPPAGSAPPITSNSSSTNPAA SSSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS			1				QATSLSGNGSWTDPVFFYVQAKRYENFIHLII
GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMVYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFILNEASVMKEFNCHH VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGEIADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGG FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVPTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVMBGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPDHISGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQURKLIKDGLIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLRRYRES/KRYR ESKKDRHMYHSLYLKVKGNVFKNKRILMEH HHLIKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEETTKK 822 2172 A 6715 772 21 DFRFGLLLPRKKKMFGFHKPKMYRSIEGCCCI SGAKSSSSRFTDSKRYEKUDFQISCFGLHETRI SGDIICNAICVLLUKRWKKLPAGSKKINWNH VVDARAGPSLKTTLKPKKVKTLISAGSRIKIST QISKLQKEFKRWINSDAHSTTSISASPAQSPLF TVNQFRWTGSDTGVGFFGSNINHPVFSFLDLI TYWKRQKICGGNYKGRFGEVLIDTHLFKPCC SNKKAIAAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTFPAAAQGQATFG NAGPLAPNGSAAPPAGSAPPITSNSSSTNPAA SSSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	1	1	ł	}		}	ALPVAVLLIVGGLVIMLYVFHRKRNNSRLGN
MSRELGQGSFGMVYEGVAKGVVKDEPETRV AIKTYNEAASMRERIEFINEASVMKEFNCHH VVRLLGVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGEIADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNEKALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRILLRRYRESKRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HKLKADKARKKLLADQAEARRSKTKEARK REEERLQAKKEEIIKTLSKEEETTK 822 2172 A 6715 772 21 DFRPGLLLPRKKKMFGFHKPKMYRSIEGCCI SGAKSSSSRFTDSKRYBKUDFQISCFGLHETR SGDICNAICVLL\LRWKKLPAGSKKINWNH VVDARAGPSILKTTLKPKVKTLSGNRIKIST QISKLQKEFKRHNSDAHSTTSISASPAQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC SNKKANAAEKPEEQGPEPLPISTQEWVTEVFM TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC SNKKANAAEKPEEQGPEPLPISTQEWVTEVFM SGDICNAICVLL\LRWKGLPGASAPNPTSNSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS			1				
AIKTVNEÄASMRERIEFLNEASVMKEFNCHH VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGEIADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARKGRHMGIGKRGTAN ARMPEKVTWMRRMRILRRLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HKLKADKARKKLLADQAEARRSKTKEARK REERLQAKKEEIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKMFGFHKPKMYRSIEGCCCI SGAKSSSSSWFTDSKRYEKIDFQSSCFGLHETRI SGDICNAICVLLUKRWKKLPAGSKKINWHH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTSISASPAQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYGTPPAAAQGQATPG SSAASGSSVPPVSSSASAPGISQISTTSSSSTNPAA SSSAASGSSVPPVSSSASAPGISQISTTSSSSTNPAA			1			ļ	l
VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGEIADGM AYLNANKFVHRDLAARNCMVAEDFITVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHISGHKAENOFOGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHYFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HIKLKADKARKKLLADQAEARRSKTKEARK REERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKKMFGFHKPKMYRSIEGC/CI SGAKSSSSIRTDSKRYEK/DFQ/SSCFGLHETR/ SGDI/CNA/CVLLUKRWKKLPAGSKK/NWMH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKRVINSDAHSITTSSASPAQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL/ TYWKRQKICCG/NYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTFPAAAGQQATPG NAGPLAPNGSSAAPPAGSSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	1	Į	1	}		1	`
SLRPEMENNPVLAPPSLSKMIQMAGEIADGM AYLMANKFVHRDLAARNCMVAEDFITVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFITYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFV\MEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENOPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMILIRLLRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HIKLKAADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKMFGPHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKYYEK\DP\OSCOP\GSFGSHETR\ SGDICNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTL\KPKVKTL\SGNRIK\ST QISKL\QKEFK\HNSDAHSTTS\SASPAQSPLL\ TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPFKYQTFPAAAGGAFTPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	Ì	1	l			Į	
AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLFVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKGDLIIRKFVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRILRRYRESSKRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HIKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRFGLLLPPKKKMFGFHKPKMYRSIEGCCI SGAKSSSSIFTDSKRYEKUDFQISCFGLHETRI SGDICNA/CVLL/LKRWKKLPAGSKK/NWNH VVDARAGPSLKTTLKPKKVKTLNSGNRIK/ST QISKLQKEFKR/HNSDAHSTTS/SASPAQSPLF TVNQFRWTGSDTGVGFFGSNRNHIFVFSFLDL/ TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC SNKKAAAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	ì						1
FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCKNTLARRKGRHMGIGKRKGTAN ARMPEK VTWMRRMILIRRLRRYRESKRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HKKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPPKKKMFGFHKPKMYRSIEGCVCI SGAKSSSSIRFTDSKRYEKUDFQISCFGLHETRI SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPSLKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASPAQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAAPPGTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	1			!			1
KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ GLSNEQVLRFVMEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKRYEK\DFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASPAQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWYTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS							
GLSNEQVLRFV/MEGGLLDKPDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSLPLPDRHSGHKAENOFOPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HKLKADKARKKLLADQAEARRSKTKEARK REERLQAKKEEIIKTLSKEEETKK DFPGGLLPRKKKMFGFHKPKMYRSIEGC/CI SGAKSSSS/RFTDSKRYEK/DFQ/SCFGLHETR/SGDICNA/CVLL/LKRWKKLPAGSKK/NWNH VVDARAGPS/LKTTLKPKKVKTL/SGNRIK/ST QISKLQKEFKR/HNSDAHSTTS/SASPAQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL/TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC SNKKAI/AAEKPEEQGPEPLPISTQEWYTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAGQQATPPA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS			1				
MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLRRYRES/KRYE ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSSSIRFTDSKRYEKIDFQ\SCFGLHETR\ SGD\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCG\NYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQQQATPQ NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS			1	1			
VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLPRKKMFGFHKPKMYRSIEGC\CI SGAKSSSSIFTDSKRYEKUDFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS]	1	i		GLSNEQVLRFV\MEGGLLDKPDNCPDMLFEL
VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLPRKKMFGFHKPKMYRSIEGC\CI SGAKSSSSIFTDSKRYEKUDFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	1	1	1	1			MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE
ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC 821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HHKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKRYEK\DFQ\\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKKLPAGSKK\\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGI\YKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS		1	1				
B21 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH HKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKRYEK\DFQ\SCFGLHETR\SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TV\QFRWTGSDTGVGFPGS\RNNHPVFSFLDL\TYWKRQKICCG\NYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTS\NSSST\PAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS			1				
821 2171 A 6691 106 825 GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH IHKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKRYEK\DFQ\SCFGLHETR\SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HINSDAHSTTS\SASPAQSPLF TVNQFR\WTGSDTGVGFPGSNRNHPVFSFLDL\TY\WKRQKICCG\NYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM SSSAKSASASAGSSVPPVSSSASAPGISQISTTSSSGFSGS		1	1			}	
SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH IHKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSS\RFTDSKRYEK\DFQ\SCFGLHETR\ SGD\\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCG\NYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	821	2171	A	6691	106	825	<u> </u>
LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH IHKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKRYEK\DFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCG\NYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	021	41/1	\ ^	10071	100	023	
VHSRARCRKNTLARRKGRHMGIGKRKGTAN ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH IHKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKRYEK\DFQ\SCFGLHETR\ SGD\\CNA\CVLL\LKRWKKLPAGSKK\\NWNH VVDARAGPS\LKTTLKYKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGS\RNHPVFSFLDL\ TYWKRQKICCG\NYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTS\NSSST\PAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS		l	l	}		}	
ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH IHKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKRYEK\DFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCG\NYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	}	1	i				
ESKKIDRIMYHSLYLKVKGNVFKNKRILMEH HKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKR\YEK\DFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCG\NYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS			l				
B22 2172 A 6715 772 21 DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSS\RFTDSKRYEK\DFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCG\NYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM B23 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS		l	l				
RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKR YEK\DFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGI\YKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS							
RREERLQAKKEEIIKTLSKEEETKK 822 2172 A 6715 772 21 DFRPGLLLPRKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKR YEK\DFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGI\YKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	İ						IHKLKADKARKKLLADQAEARRSKTKEARK
822 2172 A 6715 772 21 DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSS\RFTDSKR YEK\DFQ\SCFGLHETR\ SGD\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\H\NSDAHSTTS\SASP\AQSPLF TV\NQFR\WTGSDTGVGFPGS\NR\HPVFSFLDL\ TY\WKRQKICCG\NYKGRFGEVLIDTHLFKPCC S\NKKA\AAEKPEEQGPEPLPISTQE\WVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDS\SLLIPPKYQTPPAAAQGQATPG \NAGPLAP\NGS\AAPPAGS\AF\PTS\NSSST\PAA S\SSASGS\SV\PV\SSSA\SA\RGISQISTT\SS\SGF\SG\S]	1	}				3
SGAKSSSS\RFTDSKRYEK\DFQ\SCFGLHETR\ SGD\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\H\NSDAHSTTS\SASP\AQSPLF TV\NQFR\WTGSDTGVGFPGS\NR\HPVFSFLDL\ TY\WKRQKICCG\N\YKGRFGEVLIDTHLFKPCC S\NKKA\AA\EKPEEQGPEPLPISTQE\WVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAP\NGSAAPPAGSAF\NPTS\NSSST\NPAA SSSASGSS\VPV\SSSASAPGISQISTTSSSGFSGS	822	2172	A	6715	772	21	
SGDI/CNA/CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HINSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGI\YKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	024	1 21.72	1 ^	0,15	''-		1
VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGNIYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS			1				1 1
QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGN\YKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	1	1	1				
TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCGNYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	1		1				
TYWKRQKICCGNYKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM 823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	1						
823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	1		ٔ ا		l		
823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	}] .] .	Į ,			
823 2173 A 6727 3 4063 PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	1						
NAGPLAPNGSAAPPAGSAFNPTSNSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	823	2173	A	6727	3	4063	
SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS	ر ا	2113	ļ ^•	0/2/	~	7005	
	1						
	L	L	<u></u>			L	VOOQNESTOOISADKTQGNIGCGGDTDEGQS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						SSQPSQDGQESNVPSVGSLADPDYLNTPQMN TPVTLNSAAPASNSGAGVLPSPATPRFSVPTP RTPRTPRTPRGGGTASGQGSVKYDSTDQGSP ASTPSTTRPLNSVEPATMQPIPEAHSLYVTLIL SDSVMNIFKDRNFDSCCICACNMNIKGADVG LYIPDSSNEDQYRCTCGFSAIMNRKLGYNSGL FLEDELDIFGKNSDIGQAAERRLM/MCQSTFL PQVEGTKKPQEPPISLLLLLQNQHTQPFASLN FLDYISSNNRQTLPCVSWSYDRVQADNNDY WTECFNALEQGRQYVDNPTGGKVDEALVRS ATVHSWPHSNVLDISMLSSQDVVRMLLSLQP FLQDAIQKKRTGRTWENIQHVQGPLTWQQFH KMAGRGTYGSEESPEPLPIPTLLVGYDKDFLT ISPFSLPFWERLLLDPYGGHRDVAYIVVCPEN EALLEGAKTFFRDLSAVYEMCRLGQHKPICK VLRDGIMRVGKTVAQKLTDELVSEWFNQPW SGEENDNHSRLKLYAQVCRHHLAPYLATLQL DSSLLIPPKYQTPPAAAQGQATPGNAGPLAPN GSAAPPAGSAFNPTSNSSSTNPAASSSASGSSV PPVSSSASAPGISQISTTSSSGFSGSVGGQNPST GGISADRTQGNIGCGGDTDPGQSSSQPSQDG QESVTERERIGIPTEPDSADSHAHPPAVVTYM VDPFTYAAEEDSTSGNFWLLSLMRCYTEMLD NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY IQYLKSMAFSVYCQCRRPLPTQIHIKSLTGFGP AASIEMTLKNPERPSPIQLYSPPFILAPIKDKQT ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPNRSRRSKVSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL
824	2174	A	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR RLLPPSSAAAFSSHRHNILLCSRRRGGGGGG GGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPIHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESL\SILIS SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL QQHIATVPSLTHLCRLEIRSSLKSERLRSDSYIS

SEQ ID NO: of nucl-eotide sequence sequence sequence	hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QLPLPRSLHNYLLYEDVLRMYEVPELAAIQD G RIMGLFDRGVQMLLTTVGAFAAFSLMTIAVG TDYWLYSRGVCKTKSVSENETSKKNEEVMT HSGLWRTCCLEGNFKGLCKQIDHFPEDADYE ADTAEYFLRAVRASSIFPILSVILLFMGGLCIA ASEFYKTRHNIILSAGIFFVSAGLSNIIGIIVYIS ANAGDPSKSDSKKNSYSYGWSFYFGALSFIIA EMVGVLAVHMFIDRHKQLRATARAVTDYLQ ASAITRIPSYRYRYQRRSRSSRSTEPSHSRDA SPVGIKGFNTLPSTEISMYTLSRDPLKAATTPT
					ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA NRRTTPV
826 2176	A	6744	3	5177	SDDLRTGLFQDVQDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPEPRGLTLVRITTVPFNTT EDPDISTADLGDVLQDPCSLEYWDELQKVFV AFREFNLSESKVCELQLPDINLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTCDPL VTPTALAACTRVDSCFTPWFVPSLCVSFQFAH LEFHLCHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCKLLECRNVTMQSVVKPFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQAWQQNKCPEVEELVFSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPDPIIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIQEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMFITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEFI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVILLVSIHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI AGIVDQPMQNFQKTSEAQASAGHKAKGVISG

NO: of nucl- eotide seq- uence USSN location corresponding uence USSN location generate with the sequence location generate location location generate location location corresponding last amino generate location generate location location corresponding last amino generate location	=Glycine, H=Histidine,
nucl- eotide seq- seq- uence	=Glycine, H=Histidine,
eotide seq- uence USSN location corresponding to last amino members acid residue Q=Glutamine, R=As	
seq- uence 09/496 correspondi to last amino M=Methionine, N=, uence 914 ng to first acid residue Q=Glutamine, R=A	
uence 914 ng to first acid residue Q=Glutamine, R=A	
denote I ing to hist denotes the original interpretation Q original interpretati	
amino acid of peptide T=Threonine, V=Va	aline, W=Tryptophan,
	known, *=Stop codon,
	le deletion, \=possible
sequence nucleotide insertion	, , ,
	IGGAAELVSQTGYGILHGA
	SD\VHADQAPNSHVKYVW
KMLQSLGRPEVH	MALDVVLVRGSGQEHEGC
LLLTSEVLFVVSV	SEDTQQQAFPVTEIDCAQD
	QPRVACDVEVDGVRERLSE
	TSCHLAPSCSSMQIPCPVVA
	YLVDPHFAQVFLSKFTMVK
NKALRKGFP	
	SPGNPGRHQGPCHRPRGTK
	AAATGLEMPSSGRALLDSP
	FCSEGEGEPLALGDCFTVN
	LSCFPHTRLGKLAVVVASY LELCDDANPVDNEYFFDRS
	TGRLHVMEQLCALSFLQEI
	CRDRYFRRKELSETLDFKK
	QDFSQGPCPTVRQKLWNIL
	VISIIFVGVSIINMALMSAEL
1 1 1 1 1 1 1	VCISWFTGEFVLRFLCVRD
i i i i i i i i i i i i i i i i i i i	LLAILPFYITLLVESLSG\SQT
TQEL\ENVGAHCE	PGCLRLLRAL\RMLKAWGR
HSTGLRSLGMTIT	**************************************
STVEYFAEQSIPD	TTFTSVPCAWWWATTSMT
)	TGKIVAFMCILSGILVLALPI
1 1 1 1 1	LKLKEAAVRQREALKKLTK
l l l l l l l l l l l l l l l l l l l	DVYARSIMEMLRLKGRER
ASTRSSGGDDFW ASTRSSGGDDFW 828 2178 A 6786 5672 1360 GTHPASSGPVPLF	
	PAAVSAATREELGEPVPFV
	NPKVRSSPSGNTQSSPKSKQ PSGNPQLDSKFSNQGKQGGS
	GGHTPKALPGPGGSMGLK
	KRERSISADSFDQRDPGTPN
	HIKSQDSQHTPHSMTPSNAT
	ATEPTPAQKTPAKVVYVFS
	LKGQVETIVSFHIQNISNNK
	ALRNDPKPLPQQPPAPANQ
DQNSSQNTRLQP	TPPIPAPAPKPAAPPRPLDRE
SPGVENKLIPSVG	SPASSTPLPPDGTGPNSTPN
	SSSADPKAPPPPPVSSGEPPT
	LEHRERSLQTLRDIQRMLFP
	PQQNPGVLDGPQKKPEGPI
	KGPGPRTDVGAPFGPQGHR
	SMNSQSGTIGPDHLDHMTP
	YEEKRRKPEQVVVQQCSLQ
	VVRGPPPPYQMTPSEGWAP
	PHSLPPRGMAPHPNMPGSQ EMEGPNVPNPASRPGLSGV
	EMEGPNYPNPASKPGLSGV RNFPPGOGIFSGPGRGERFP
	QLAEKQLGLPPGMAMEGIR
	QRHMEPGNNPIFPRIPVEGP
	POMGPGRELEFGMVPSGM
	SNSQMIPQKMREAGAGPEE
	LPAQQKMVPLPFGEHPQQE
	QGPGSNSGLRNLREPIGPDO
	PLNPSSNPTSLNTAPPVQRG
	SQVHSPGINPLKSPTMHQVQ
SPMLGSPSGNLKS	SPQTPSQLAGMLAGPAAAA
SIKSPPVLGSAAA	SPVHLKSPSLPAPSPGWTSS
PEPPLQSPGIPPNI	HKAPLTMASPAMLGNVESG
GPPPPTASQPASV	NIPG\SLPSSTPYTMPPEPTL
SQNPLSIMMSRW	MSKFAM\PS\SNPGYNHDAI

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KTVASSDDDSPPARSPNLPSMNNMPGMGINT QNPRISGPNPVVPMPTLSPMGMTQPLSHSNQ MPSPNAVGPNIPPHGVPMGPGLMSHNPIMGH GSQEPPMVPQGRMGFPQGFPPVQSPPQQVPFP HNGPSGGQGSFPGGMGFPGEGPLGRPSNLPQ SSADAALCKPGGPGGPDSFTVLGNSMPSVFT DPDLQEVIRPGATGIPEFDLSRIIPSEKPSQTLQ YFPRGEVPGRKQPQGPGPGFSHMQGMMGEQ APRMGLALPGMGGPGPVGTPDIPLGTAPSMP GHNPMRPPAFLQQGMMGPHHRMMSPAQST MPGQPTLMSNPAAAVGMIPGKDRGPAGLYT
829	2179	A	6797	433	3	HPGPVGSPGMMMSMQGMMGP\NRTS ASFFNFSICICKIILEVGPPVGHPAHDDVGGRH GPGGR/GSRSPRSLQCAPGGGRRSGCPAGSSP ASTCPPSPGGSGADRFGPSPPPPSREAAPTAG AAASSTSSGASCPPVPASSRWGVRSRTRSGSG GEREPRDRPSERPRLV
830	2180	A	6800	3 _	1911	LPERAFGPRTPRAPRRRRRRLLLSPPPRPPPPL DREPRAPGPWLCPSRAGTAQDPARIRERRGR VAGGAAGPAMELRARGWWLLCAAAALVAC ARGDPASKSRSCGEVRQIYGAKGFSSS\DVPQ AEISGEHLRICPQGYTCCTSEMEENLANRSHA ELETALRDSSRVLQAMLATQLRSFDDHFQHL LNDSERTLQATFPGAFGELYTQNARAFRDLY SELRLYYRGANLHLEETLAEFWARLLERLFK QLHPQLLLPDDYLDCLGKQAEALRPF\GEAP\ RELRLATRA\FVAAR\SFVQGLGVAS\DVVR KVAQVPLG\PEC\SRAVIEAGSYC\ALHCVGVP GARPCPDYCRNVLKGCLANQADLDAEWRNL LDSMVLITDKFWGTSGVESVIGSVHTWLAEA INALQDNRDTILTAKVIQGCGNPKVNPQGPGP EEKRRGKLAPREPPSGTLEKLVSEAKAQL RDVQDFWISLPGTLCSEKMALSTASDDRCWN GMARGRYLPEVMGDGLANQINNPEVEVDIT KPDMTIRQQIMQLKIMTNRLRSAYNGNDVDF QDASDDGSGSGSGDGCLDDLCGRKVSRKSSS SRTPLTHALPGLSEQEGQKTSAASCPQPPTFL LPLLLFLALTVARPRWR
831	2181	A	6808	2	1522	ASRHGMTPGALLMLLGALGPPLAPGVRGSEA EGRLREKLFSGYDSSVRPAREVGDRVRVSVG LILAQLISLNEKDEEMSTKVYLDLEWTDYRLS WDPAEHDGIDSLRITAESVWLPDVVLLNNND GNFDVALDISVVVSSDGSVRWQPPGIYRSSCS IQVTYFPFDWQNCTMVFSSYSYDSSEVSLQT GLGPDGQGHQEIHIHEGTFIENGQWENIHKPS RLIQPPGDPRGGREGQRQEVIFYLIIRKPLFY LVNVIAPCILITLLAIFVFYLPDAGEKMGLSIF ALLTLTVFLLLLADKVPETSLSVPIIIKYLMFT MVLVTFSVILSVVVLNLHHRSPHTHQMPLWV RQIFIHKLPLYLRLKRPKPERDLMPEPPHCSSP GSGWGRGTDEYFIRKPPSDFLFPKPNRFQPEL SAPDLRRFIDGPNRAVALLPELREVVSSISYIA RQLQEQEDHDALKEDWQFVAMVVDRLFLW TFIIFTSVGTL\VIFLDATYHLPPPDPFP
832	2182	A	6824	71	1079	ETMAKNPPENCEDCHILNAEAFKSKKICKSLK ICGLVFGILALTLIVLFWGSKHFWPEVPKKAY DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGN GTDETLEVHDFKNGYTGIYFVGLQKCFIKTQI KVIPEFSEPEEEIDENEEITTTFFEQSVIWVPAE KPIENRDFLKNSKILEICDNVTMYW\INPTL\IS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTFAKQLHHNFAFIILVSELQDFEEEGEDLHFP ANEKKGIEQNEQWVVPQVKVEKTRHARQAS
833	2183	A	6846	116	602	EEELPINDYTENGIEFDPMLDERGYCCIYCRR GNRYCRRVCEPLLGYYPYPYCYQGGRVICRV IMPCNWWVARMLGRV EAEGEQVCGAKCCGDAPHVENREEETARIGP
						GVMESKEERALNNLIVENVNQENDEKDEKE QVANKGEPLALPLNVSEYCVPRGNRRRFRVR QPILQYRWDIMHRLGEPQARMREENMERIGE EVRQLMEKLREKQLSHSLRAVSTDPPHHDHH DEFC\LMP
834	2184	A	6851	3	2024	PNGVALLHLPGAAVIPNTNYMFQDALGGRSR GSREESPAPSRAPASASLWRRLVVVEAKMAA HAAAAAQAAAAQAAHAEAADSWYLALLGF AEHFRTSSPPKIRLCVHCLQAVFPFKPPQRIEA RTHLQLGSVLYHHTKNSEQARSHLEKAWLIS QQIPQFEDVKFEAASLLSELYCQENSVDAAKP LLRKAIQISQQTPYWHCRLLFQLAQLHTLEKD LVSACDLLGVGAEYARVVGSEYTRALFLLSK GMLLLMERKLQEVHPLLTLCGQIVENWQGN PIQKESLRVFFLVLQVTHYLDAGQVKSVKPC LKQLQQCIQTISTLHDDEILPSNPADLFHWLP KEHMCVLVYLVTVMHSMQAGYLEKAQKYT DKALMQLEKLKMLDCSPILSSFQVILLEHIIM CRLVTGHKATALQEISQVCQLCQQSPRLFSN HAAQLHTLLGLYCVSVNCMDNAEAQFTTAL RLTNHQELWAFIVTNLASVYIREGNRHQEVVI LYSLLERINPDHSFPVSSHCLRAAAFYVRGLF SFFQGRYNEAKRFLRETLKMSNAEDLNRLTA CSLVLLGHIFYVLGNHRESNNMVVPAMQLAS KIPDMSVQLWSSALLRDLNKACGNAMDAHE AAQMHQNFSQQLLQDHIEACSLPEHNLITWT DGPPPVQFQAQNGPNTSLASLL
835	2185	A	6862	334	1268	PTRRPILPLTSPKAISVPSPLQGKQHTLVKSCL SVSGIGGFLVSLSSRMKLQTLAVSVTALKFWS AYVPCQTQDRDALRLTLEQIDLIRRMCASYSE LELVTSAKALNDTQKLACLIGVEGGHSLDNS LSILRTFYMLGVRYLTLTHTCNTPWAESSAK GVHSFYNNISGLTDFGEKVVAEMNRLGMMV DLSHVSDAVARRALEVSQAPVIFSHSAARGV CNSARNVPDDILQLLEEERWAFVMVSLFHGE LIQWQPIRPMCSTVADHFDHIKAVIGSKFIGI GGDYDGAGKYRKKTTCKAPWRTSSRMSS
	2180		0802	313	11	PPRSRPSCWRKKVGPGRPWWWGGTGPPGQG RPEIRLLPLPMTGACGAVAASRTGSSGPG/SSL PNGHGGKGSGLANGLAGNP\GHLGLGSSFGT GPGSGRPPP
837	2187	A	6863	2	1615	VLRGQRGPAGGLAEERRRGRNEWRIHDVTT APFPGLVQRRSRLLIVSQVRYFLKNKVSPDLC NEDGLTALHQCCIDNFEEIVKLLLSHGANVN AKDNELWTPLHAAATCGHINLVKILVQYGA DLLAVNSDGNMPYDLCEDEPTLDVIETCMAY QGITQEKINEMRVAPEQQMIADIHCMIAAGQ DLDWIDAQGATLLHIAGANGYLRAAELLLDH GVRVDVKDWDGWEPLHAAAFWGQMQMAE LLVSHGANLNARTSMDEMPIDLCEEEFKVL LLELK\HKHDVIMKSQLRHKSSLSRRTSHRQA S/SVGKVVRRTQPVGTGPNL\YRKEYE/GEEAI LWQRSA\AEDQRTSTYNGDIRET\RTDQENKD

						(A. Ali C. Ci
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	1	in	nucleotide	location	
eotide	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ł	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	İ			peptide		/=possible nucleotide deletion, \=possible
				sequence	L	nucleotide insertion
						PNPRLEK\PVLLSEFPTKIPRGELDMPVENGLR
	ļ	İ		į	ļ	APVSAYQYALANGDVWKVHEVPDYSMAYG
	Ì			j		NPGVADATPPWSSYKEQSPQTLLELKRQRAA
	l .					AKLLSHPFLSTHLGSSMARTGESSSEGKAPLI
	1				!	GGRTSPYSSNGTSVYYTVTSGDPPLLKFKAPI
						EEMEEKVHGCCRIS
838	2188	A	6865	6291	739	AGPLEPRVQGAMALQLWALTLLGLLGAGAS
4		1			[LRPRKLDFFRSEKELNHLAVDEASGVVYLGA
	İ		1		[VNALYQLDAKLQLEQQVATGPVLDNKKCTP
		İ				PIEASOCHEAEMTDNVNQLLLVDPPRKRLVE
						CGQLLKGI\CALRALSNISLRLFYEDGSGEKSF
		1		1		VASNDEGVATVGLVSSTGPGGDRVLFVGKG
	1]		1	1	NGPHDNGIIVSTRLLDRTDSREAFEAYTDHAT
				[1	YKAGYLSTNTQQFVAAFEDGPYVFFVFNQQD
				1		KHPARNRTLLARMCREDPNYYSYLEMDLQC
				1	1	RDPDIHAAAFGTCLAASVAAPGSGRVLYAVF
	1	1	}			SRDSRSSGGPGAGLCLFPLDEVHAKMEANRN
		i	İ			ACYTGTREARDIFYKPFHGDIQCGGHAPGSSK
		i		l	1	SFPCGSEHLPYPLGSRDGLRGTAVLQRGGLN
		}			1	
						LTAVTVAAENNHTVAFLGTSDGRILKVYLTP
	1	1			ļ	DGTSSEYDSILVEINKRVKRDLVLSGDLGSLY
		Ì	1	l	1 .	AMTQDKVFRLPVQECLSYPTCTQCRDSQDPY
		1				CGWCVVEGRCTRKAECPRAEEASHWLWSRS
		1				KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA
	1		1	1	İ	LSEEDELLCLFGESPPHPARVEGEAVICNSPSS
	1	1				IPVTPPGQDHVAVTIQLLLRRGNIFLTSYQYPF
	i	1	1		}	YDCRQAMSLEENLPCISCVSNRWTCQWDLR
	ļ			ļ		YHECREASPNPEDGIVRAHMEDSCPQFLGPSP
	1	1	l	1		LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD
]	1	LLKFMEPVTMQESGTFAFRTPKLSHDANETL
						PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC
	!					SLCRAANPDYRCAWCGGQSRCVYEALCNTT
	ł	1	i	•		SECPPPVITRIQPETGPLGGGIRITILGSNLGVQ
			ļ	1		AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA
	İ					AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP
İ						KPLSVEPQQGPQAGGTTLTIHGTHLDTGSQED
,			1	ļ		VRVTLNGVPCKVTKFGAQLQCVTGPQATRG
ļ	1	1		1		QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE
ļ	1			ļ		PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP
		1	1			LQSWQPPREAESLQPMTVVGTDYVFHNDTK
1	i					VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT
1	1			İ		EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR
		1				GTNLNKAMTLQEAEAFVGAERCTMKTLTET
1	1		1			DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF
						GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM
l		1				VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG
1		1	1	Į		LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG
l	1		1	ĺ		IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL
	1			1		DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT
				1	1	LIENQPEFSARAKVYFASLLTVALHGKLEYYT
1		1	1	1	1	DIMHTLFLELLEQYVVAKNPKLMLRRSETVV
1		1	1	Į.		ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI
1			Į.	1		KHQVEKGPVDAVQKKAKYTLNDTGLLGDD
		1		1		VEYAPLTVSVIVODEGVDAIPVKVLNCDTISQ
		1	1	1		
	1	1	1			VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG
1	1		4	1	l .	STAQILSDLDLTSQREGRWKRVNTLMHYNVR
l .					1	DO 100 H 07/1/01/00 0000000000000000000000000000
						DGATLILSKVGVSQQPEDSQQDLPGERHALL
						EEENRVWHLVRPTDEVDEGKSKRGSVKEKE

COTO ID	CEO ID	Mat	OT C	Dendistad	Dundint - 1 1	L Amino poid opposed A A Latin C. C. C.
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	İ		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		,		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł	}	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i		ļ		peptide	"	/=possible nucleotide deletion, \=possible
		ļ		sequence		nucleotide insertion
						HIWKTNSLPLRFWVNILKNPHFIFDVHVHEVV
l	1			1	{	DASLSVIAQTFMDACTRTEHKLSRDSPSNKLL
	1			ł	1	VANCIOTATION OF DATA CIRCLE CONTROLLE
	· .				j	YAKEISTYKKMVEDYYKGIRQMVQVSDQDM
		l				NTHLAEISRAHTDSLNTLVALHQLYQYTQKY
						YDEINALEEDPAAQKMQLAFRLQQIAAALE
					ļ	NKVTDL
839	2189	A	6872	1	1485	RARRLALQCHVCVCALTPGEQSGRRLPGQT
					ŀ	WLMFSCFCFSLQDNSFSSTTVTECDEDPVSLH
	ł	İ				EDQTDCSSLRDENNKENYPDAGALVEEHAPP
ł	l	·			i	SWEPQQQNVEATVLVDSVLRPSMGNFKSRKP
						KSIFKAESGRSHGESQETEHVVSSQSECOVRA
	ľ					GTPAHESPQNNAFKCQET\VRL\QPRIDQRTAT
İ	Ì	ł				
1		[ĺ	SPKDAFETR\QDLNEEEAAQVHGVKDPAPAS
				ļ		TQSVLA\DGTDSADPSPVHKDGQNEADSAPE
	[DLHSVGTSRLLL/YHITDGDNPTAVRHGCSL/F
[Į				1	SGQSQRFNLDPESAPSPPSTQQFMMPRSSSRC
	1	l				SCGDGKEPQTITQLTKHIQSLKRKIRKFEEKFE
		1		1		QEKKYRPSHGDKTSNPEVLKWMNDLAKGRK
	l	İ				QLKELKLKLSEEQGSAPKGPPRNLLCEQPTVP
l	1	1				RENGKPEAAGPEPSSSGEETPDAALTCLKERR
[1					EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL
	1	l .			j	STPSLIPTIVSQDTCMLLLCTDV
840	2190	A	6873	2	2054	FFRFYFSFIRLFAMSLADLTKTNIDEHFFGVAL
0.00	2170	^`	0075		2034	
	1					ENNRRSAACKRSPGTGDFSRNSNASNKSVDY
l	ł	ł	1			SRSQCSCGSLSSQYDYSEDFLCDCSEKAINRN
						YLKQPVVKEKEKKKYNVSKISQSKGQKEISV
		l				EKKHTWNASLFNSQIHMIAQRRDAMAHRILS
		1				ARLHKIKGLKNELADMHHKLEAILTENQFLK
		1			1	QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV
	1	- '				KNLRQLLRKSQEKERTLSRKLRETDSQLLKT
ì						KDILQALQKLSEDKNLAEREELTHKLSIITTK
		}				MDANDKKIQSLEKQLRLNCRAFSRQLAIETR
·	(ľ	ĺ		1	KTLAAQTATKTLQVEVKHLQQKLKEKDREL
						EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD
				,		RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG
	ŀ					
				İ		NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY
				·		EDLSGEEKHLEVQILLENTGRQKDKKEDQEK
	1					KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR
1	1					EKFKRSMQRNGVDDT\LGKGTAPYTKGPLRQ
		İ				RRHYSFTEATENLHHGLPASGGPANAGNMR
•	l	1				YSHSTGKHLSNREEMELEHS\DSGYEPSFGKS
						SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD
	[l				QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA
	l ·	1				FGDSKVTVVNSIKPSSPTEGKRKIII
841	2191	A	6874	3	2867	SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG
	l		55, 1	-	200,	NAPAPGTPAASGWQPPTYHSGRAFSARYPRP
		1		į		SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA
}	{	}		}		DHAVRPLHGARGGQPPVPQQHVLERQVQLS
		1				QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE
		l				DTPWSDQRPREGEGEPPRGQLQPSRPTRARG
		[}	TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP
	1			}		REPRRTVSESVIAVKASFPSSALPPRTGVALG
						RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP
·		1				SGSVGGPARPASGPRQAREASLVVTCRTNKF
	l	ł				RKNNYKWVAASSKSPRVARRALSPRVAAEN
	ľ	ľ				VCKASAGMANKVEKPQLIADPEPKPRKPATS
!	}				1	SKPGSAPSKYKWKASSPSASSSSFRWQSEAG
	[1			,	SKDHASQLSPVLSRSPSGD\RPALAHSGLKPLS
						GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG
l	j					TSPAATAKSHLSLRRRQALRGKSSPVLKKTPN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						VRTAPTSKVIKTRYRIVKKTPASPLSAPPFPLS LPSWRARRLSLSRSLVLNRLRPVASGGKAQ PGSPWWRSKGYRCIGGVLYKVSANKLSKTSG QPSDAGSRPLLRTGRLDPAGSCSRSLASRAVQ RSLAIIRQARQRREKRKEYCMYYNRFGRCNR GERCPYIHDPEKVAVCTRFVRGTCKKTDGTC PFSHHVSKEKMPVCSYFLKGICSNSNCPYSHV YVSRKAEVCSDFLKGYCPLGAKCKKKHTLLC PDFARRGACPRGAQCQLLHRTQKRHSRRAAT SPAPGPSDATARSRVSASHGPRKPSASQRPTR QTPSSAALTAAAVAAPPHCPGGSASPSSSKAS SSSSSSSSPPASLDHEVAPSLQEAALAAACSNR LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKDS GKPLHIKPRL
842	2192		6898	506	2071	WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH RNKFKVINVDDDGNELGSGIMELTDTELILYT RKRDSVKWHYLCLRRYGYDSNLFSFESGRRC QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFAA QNLPNGYPRYPSFGDASSHPSSRHPSVGSARL PSVGEESTHPLLVAEEQVHTYVNTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVNKLVYENINGLSIPSASGVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTPTRRTELYAVIDIERTAAMSNLQKAL PRDDGTSRKTRHNSTVDLPL
843	2193	A	6919	2	663	AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS L/VFLGPGLYNN/GSSMCGAE\EPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF
844	2194	A	6928	902	366	HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM
845	2195	A	6939	1660	317	LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRRIWRILEEKESVAGAVQTLLLRSQE GGV1TSAAASTLSEPPRRTQESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHLVPHFKPWLVHPEQSP LPSLALSELSVQHADSVLENIDESAVAESREE RVMGGAGGEGVSDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP RRVLQPGWTEPQ

SEQ ID	SEQ ID	Met	T CEO	Prodicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	SEQ ID NO:	Predicted beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ļ				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
846	2196	Α	6944	42	2672	RRKMAGCRGSLCCCCRWCCCCGERETRTPE
		ĺ	<u> </u>		ĺ	ELTILGETQEEEDEILPRKDYESLDYDRCINDP
i						YLEVLETMDNKKGRRYEAVKWMVVFAIGV
·		!		1		CTGLVGLFVDFFVRLFTQLKFGVVQTSVEECS
		1				QKGCLALSLLELLGFNLTFVFLESLLGLIEPVE
		1				AGSGITEGKCYLYARQVPGLVRLPTLLWKAL
					į	GVLLTVAAMLLI\GLGSPMIHSGSVVGAGLPQ
		1			1	FQSISLRKIQFNFPYFRSDRYGK\DKRDFVSAG AAAGVAAAFGAPIGGTLFSLEEGSSFWNQGL
						TWKVLFCSMSATFTLNFFRSGIQFGSWGSFQL
						PGLLNFGEFKCSDSDKKCHLWTAMDLGFFV
		ŀ				VMGVIGGLLGATFNCLNKRLAKYRMRNVHP
		ļ			· ·	KPKLVRVLESLLVSLVTTVVVFVASMVLGEC
		İ				RQMSSSSQIGNDSFQLQVTEDVNSSIKTFFCP
				ļ		NDTYNDMATLFFNPQESAILQLFHQDGTFSPV
		1				TLALFFVLYFLLACWTYGISVPSGLFVPSLLC
		}			}	GAAFGRLVANVLKSYIGLGHIYSGTFALIGAA
		ļ	Ì			AFLGGVVRMTISLTVILIEST\NEITYGLPIMVT
						LMVGKWTGDFFNKGI\YDIHVGLRGVPLLEW
	0					ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV
]			SILRTTVHHAFPVVTENRGNEKEFMKGNQLIS
						NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL
]			·	RNMCDEHIASEEPAEKEDLLQQMLERRYTPY
		1				PNLYPDQSPSEDWTMEERFRPLTFHGLILRSQ
		ŀ				LVTLLVRGVCYSESQSSASQPRLSYAEMAED
. 1		1				YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE
						IVGITRHNLTYEFLQARLRQHYQTI
847	2197	A	6951	3	1994	NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER
•		'-	0,51	1	1,551	LSSIEKIKQLREQVNDLFSRKFGEAIGVDFPVK
						VPYRKITFNPGCVVIDGMPPGVVFKAPGYLEI
		1			·	SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST
ľ	•		İ			VGKRKIDQEGRVFQEKWERAYFFVEVONIST
						CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY
		ľ	l			MERMRDEKLHELKKGLRKYLLGLSDTECPE
		[QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR
		1	}	}		EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE
		1				NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK
		Ì				NFCINWSKLVSVASTGTPPMVDANNGLVTKL
				·		KSRVATFCKGAELKSICCIIHPESLCAQ\KLKM
	}	l	}	}		DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL
		Ī				DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI
ļ		}				DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL
ļ]		,		WETHLTRNNLAHFPTLKLVSRNESDGLNYIP
ĺ			1	!		KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI
						DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE
		1				FYKYLWGSYPKYKHHCAKILSMFGSTYICEO
		l	1		Ì	LFSIMKLSKTKYCSQLKDSQWDSVLHIAT
848	2198	A	6985	3	289	SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK
· · · · · · · · · · · · · · · · · · ·		l		-		ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C
l		1	}	ľ		CFNAINTKIPIQRLESYTRITNIQCPKEAVM
			6000	963	5	LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM
849	2199	A	6999			, and the second
849	2199	A	6999			LLFGLFSLFYVFTLLGNGTILGLISLDSRIHAP
849	2199	A	6999			LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP
849	2199	A	6999			MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP
849	2199	A	6999	,		MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT
849	2199	A	6999			MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA
849	2199	A	6999			MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GLFYGTAIIMYVGPRYGNPKEQKKYLLLFHS LFNPMLNPLICSLRNSEVKNTLKRVLGVERAL MGNDSVSYEYGDYSDLSDRPVDCLDGACLAI
						DPLRVAPLPLYAAIFLVGVPGNAMVAWVAG KVARRRVGATWILHLAVADILICCLSLPILAV PIARGGHWPYGAVGCRALPSIILLTMYASVLL LAALSADLCFLALGPAWCLRFS/GACGVQVA CGAAWTLALLLTVPSAIYRRLHQEHFPARLQ CVVDYGGSSSTENAVTAIRFLFGFLGPLVAVA SCHSALLCWAARRCRPLGTAIVVGFFVCWAP YHLIGLVLTVAAPNSALLARALRAEPLIVGL ALAHSCLNPMLFLYFGRAQLRRSLPAACHW ALRESQGQDESVDSKKSTSHDLVSEMEV
851	2201	A	7011		2310	AAASPLRMSRKGPRAEVCADCSAPDPGWASI SRGVLVCDECCSVHRSLGRHISIVKHLRHSA WPPTLLQMVHTLASNGANSIWEHSLLDPAQV QSGPALKQTPKDKV\HPIKSEFIRAKYQMLAF VHKLPCRDDDGVTAKDLSKQLHSSVRTGNLE TCLRLLSLGAQANFFHPEKGTTPLHVAAKAG QTLQAELLVVYGADPGSPDVNGRTPIDYARQ AGHHELAERLVECQYELTDRLAFYLCGRKPD HKNGHYIIPQMADSLDLSELAKAAKKLQAL SNRLFEELAMDVYDEVDRRENDAVWLATQN HSTLVTERSAVPFLPVNPEYSATRNQGRQKL ARFNAREFATLIIDILSEAKRRQQGKSLSSPTD NLELSLRSQSDLDDQHDYDSVASDEDTDQEP LRSTGATRSNRARSMDSSDLSDGAVTLQEYL ELKKALATSEAKVQQLMKVNSSLSDELRRLQ REIHKLQAENLQLRQPPGPVPTPPLPSERAEH TPMAPGGSTHRRDRQAFSMYEPGSALKPFGG PPGDELTTRLQPFHSTELEDDAIYSVHVPAGL YRRKGVSASAVPFTPSSPLLSCSQEGSRHTSK LSRHGSGADSDYENTQSGDPLLGLEGKRFLE LGKEEDFHPELESLDGDLDPGLPSTEDVILKT EQVTKNIQELLRAAQEFKHDSFVPCSEKIHLA VTEMASLFPKRPALEPVRSSLRILNASAYRLQ SECRKTVPPEPGAPVDFQLLTQQVIQCAYDIA
8.52	2202	Ā	7016	484	1777	RISKIQVYYSTGYSSRKMNPTLGLAIFLAVLL TVKGLLKPSFSPRNYKALSEVQGWKQRMAA KELARQNMDLGFKLLKKLAFYNPGRNIFLSP LSISTAFSMLCLGAQDSTLDEIKQGFNFRKMP EKDLHEGFHYIIHELTQKTQDLKLSIGNTLFID QRLQPQRKFLEDAKNFYSAETILTNFQNLEM AQKQINDFIESKTHGKINNLIENIDPGTVMLL ANYIFFRARWKHEFDPNVTKEEDFFLEKNSS VKVPMMFRSGIYQVGYDDKLSCTILEIPYQK NITAIFILPDEGKLKHLEKGLQVDTFSRWKTL LSRRVVDVSVPRLHMTGTFDLKKTLSYIGVS KIFEEHGDLTKIAPHRSLKVGEAVNKAELKM DERGTEGAAGTGAQTLPMETPLVVKIDKPYL LLIYSEKIPSVLFLGKIVNPIGK
853	2203	A	7017	1	3293	MTHACNPSTLGGQGRRITRSHGRRRSSRGPV ARHVAAGAGHENKHGGSRRFPAGVAPRRAM ANVSKKVSWSGRDRDDEEAAPLLRRTARPG GGTPLLNGAGPGAARQSPRSALFRVGHMSSV ELDDELLEP\DMDPPHPFPKEIPHNEKLLSLKY ESLDYDNSENQLFLEEERRINHTAFRTVEIKR WVICALIGILTGLVACFIDIVVENLAGLKYRVI KGSILPNIDKFTEKGGLSFSLLLWATLNAAFV

CODO TO	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino soid sagues (A - Alexino C-Custoino
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	,,,,,,,	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ŀ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		İ	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
	[[amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide		/=possible nucleotide deletion, \=possible
	1		1	sequence]	nucleotide insertion
			T			LVGSVIVAFIEPVAAGSGIPQIKCFLNGVKIPH
				ļ		VVRLKTLVIKVSGVILSVVGGLAVGKEGPMI
	ł	1	l			HSGSVIAAGISQGRSTSLKRDFKIFEYFRRDTE
		ľ				KRDFVSAGAAAGVSAAFGAPVGGVLFSLEEG
						ASFWNQFLTWRIFFASMISTFTLNFVLSIYHG
						NMWDLSSPGLINFGRFDSEKMAYTIHEIPVFI
	l	}	l	i	}	AMGVVGGVLGAVFNALNYWLTMFRIRYIHR
						PCLQVIEAVLVAAVTATVAFVLIYSSRDCQPL
						QGGSM\$YPLQLFCADGEYNSMAAAFFNTPEK
						SVVSLFHDPPGSYNPLTLGLFTLVYFFLACWT
		1	1			YGLTVSAGVFIPSLLIGAAWGRLFGISLSYLTG
İ	1	ĺ				AAIWADPGKYALMGAAAQLGGIVRMTLSLT
1		l	ł			VIMMEATSNVTYGFPIMLVLMTAKIVGDVFIE
			[GLYDMHIQLQSVPFLHWEAPVTSHSLTAREV MSTPVTCLRRREKVGVIVDVLSDTASNHNGF
						PVVEHADDTOPARLOGLILRSOLIVLLKHKVF
					ļ	VERSNLGLVORRLRLKDFRDAYPRFPPIOSIH
Ì		l				VSQDERECTMDLSEFMNPSPYTVPQEASLPR
1						VFKLFRALGLRHLVVVDNRNQVVGLVTRKD
]	ļ	ļ]	j	j	LARYRLGKRGLEELSLAQTGPKAQATAEGRV
						AGAAQQPCQLRAVTLEDLGLLLAGGLASPEP
					-	LSLEELSERYESSHPTSTASVPEQDTAKHWNQ
!						LEQWVVELQAEVACLREHKORCERATRSLL
ļ		}			l	RELLQVRARVQLQGSELRQLQQEARPAAQAP
						EKEAPEFSGLQNQMQALDKRLVEVREALTRL
			-			RRRQVQQEAERRGAEQEAGLRLAKLTDLLQ
1		ľ		ĺ	Į.	QEEQGREVACGALQKNQEDSSRRVDLEVAR
						M
854	2204	A	7037	139	2604	AGTWEPRPYDQAKETGAPGSQPPVPPMELRP
			ļ			WLLWVVAATGTLVLLAADAQGQKVFTNTW
	1	[1		ĺ	AVRIPGGPAVANSVARKHGFLNLGQIFGDYY
						HFWHRGVTKRSLSPHRPRHSRLQREPQVQWL
		1		}		EQQVAKRRTKRDVYQEPTDPKFPQQWYL\SG
						VTQ\RDLMVKAAWAQGYTGHGIVVSILDDGI
						EKNHPDLAGNYDPGASFDVNDQDPDPQPRY TOMODNIBHGTBCAGEVAAVANNGVCGVGV
1	İ			ŀ		TQMNDNRHGTRCAGEVAAVANNGVCGVGV AYNARIGGVRMLDGEVTDAVEARSLGLNPN
1	1	1			ĺ	HIHIYSASWGPEDDGKTVDGPARLAEEAFFR
		1			1	GVSOGRGGLGSIFVWASGNGGREHDSCNCD
!]	j	1	1	1	GYTNSIYTLSISSATQFGNVPWYSEACSSTLA
						TTYSSGNQNEKQIVTTDLRQKCTESHTGTSAS
		ļ				APLAAGIIALTLEANKNLTWRDMQHLVVQTS
					1	KPAHLNANDWATNGVGRKVSHSYGYGLLD
	1	1		1		AGAMVALAQNWTTVAPQRKCIIDILTEPKDI
				1		GKRLEVRKTVTACLGEPNHITRLEHAQARLT
	1	1				LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY
ĺ				1	ı]
		ĺ				SADGFNDWAFMTTHSWDEDPSGEWVLEIEN
						TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS
			1			
						TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS
						TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS
						TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS
						TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS
						TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD
						TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL
855	2205	A	7058	3	1441	TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL QRPASQLLAPFAAEALPGAPRAAMAQHFSLA
855	2205	A	7058	3	1441	TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL QRPASQLLAPFAAEALPGAPRAAMAQHFSLA ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF
855	2205	A	7058	3	1441	TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL QRPASQLLAPFAAEALPGAPRAAMAQHFSLA ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF LVKEKGYDKELLNVTPEDWDFCCKGLALDL
855	2205	A	7058	3	1441	TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL QRPASQLLAPFAAEALPGAPRAAMAQHFSLA ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YFDLPGALLCARVVDYLTKLNNGQKTFDFW KDIVAAIQHNYKMSAFKENCGIYFPEIKRDPG
						RYLHSRPESVKKWLRQLKNAGKILLLITSSHS DYCRLLCA\YILGNDFTDLFDIVITNALKPGFP SHLPSQRPFRTLENDEEQEALPSLDKPGWYSQ GNAVHLYELLKKMTGKPEPKVVYFGDSMHS DIFPARHYSNWETVLILEELRGDEGTRSQRPE ESEPLEKKGKYEGPKAKPLNTSSKKWGSFF\I DSVLGLENTEDSLVYTWSCKRISTYSTIAIPSI EAIAELPLDYKFTRFSSSNSKTAGYYPNPPLV LSSDETLISK
856	2206	A	7082	396	1635	SSPSVFEFEHAVQPVFTMEFLKTCVLRRNACT AVCFWRSKVVQKPSVRRISTTSPRSTVMPAW VIDKYGKNEVLRFTQNMMMPIIHYPNEVIVK VHAASVNPIDVNMRSGYGATALNMKRDPLH VKIKGEEFPLTLGRDVSGVVMECGLDVKYFK PGDEVWAAVPPWKQGTLSEFVVVSGNEVSH KPKSLTHTQAASLPYVALTAWSAINKVGGLN DKNCTGKRVLILGASGGVGTFAIQVMKAWD AHVTAVCSQDASELVRKLGADDVIDYKSGSV EEQLKSLKPFDFILDNVGGSTETWAPDFLKK WSGATYVTLVTPFLLNMDRLGIADGMLQTG VTVGSKALKHFWKGVHYRWAFFMASGPCL DDIAELVDAGKIRPVIEQTFPFSKVPF AFLKV ERGHARGKTVINVV
857	2207	A	7088	320	2417	LRRRKMTPQSLLQTTLFLLSLLFLVQGAHGR GHREDFRFCSQRNQTHRSSLHYKPTPDLRISIE NSEEALTVHAPFPAAHPASRSFPDPRGLYHFC LYWNRHAGRLHLLYGKRDFLLSDKASSLLCF QHQEESLAQGPPLLATSVTSWWSPQNISLPSA ASFTFSFHSPPHTGAHNASVDMCELKRDLQL LSQFLKHPQKASRRPSAAPASQQLQSLESKLT SVRFMGDMGSFEEDRINATVWKLQPTAGLQ DLHIHSRQEEEQSEIMEYSVLLPRTLFQRTKG RSGEAEKRLLLVDFSSQALFQDKNSSQVLGE KVLGIVVQNTKVANLTEPVVLTFQHQLQPKN VTLQCVFWVEDPTLSSPGHWSSAGCETVRRE TQTSCFCNHLTYFAVLMVSSVEVDAVHKHY LSLLSYVGCVVSALACLVTIAAYLCSRVPLPC RRKPRDYTIK VHMNLLLAVFLLDTSFLLSEPV ALTGSEAGCRASAIFLHFSLLTCLSWMGLEG YNLYRLVVEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPIILAVHRTPEGVIYPS MCWIRDSLVSYITNLGLFSLVFLFNMAMLAT MVVQILRLRPHTQKWSHVLTLLCLSLVLGILP WALIFFSFASGTFQLVVLYLFSITTSFQGFLIFI WYWSMRLQARGGPSPLKSNSDSARLPISSGS TSSSRI
858	2208	A	7091	185	415	DAGAVKSSDTNIWFRGMCDDKKGHRCPS*G QPQHFHVAFHTEAEGAMFYFRLHVIHRVMQS QQQLFPSTLFSWLLE
859	2209	A	7136	3	302	FFFWRQSLALLPRLECSGATGAHCNLHFPGSS DCPTSAS*IAGITGACYHAWLLFVFLAETGFH HVGQGGLELLTSSDPSGSASQSAGITGVSHCT WPI
860	2210	A	7156	23	591	ALSTETRTPDMRRLLLVTSLVVVLLWEAGAV PAPKVPIKMQVKHWPSEQDPEKAWGARVVE PPEKDDQLVVLFPVQKPKLLTTEEKPRGQGR GPILPGTKAWMETEDTLGRVLSPEPDHDSLY

CEO ID	CEO TO	Mos	CEO	Predicted	Predicted end	Amino said cogner (A. Ali C. Ci
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
NO: 01 nucl-	peptide	l noa	in NO:	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	uence	1	09/496	correspondi	to last amino	
seq-	uence ,		914		acid residue	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first amino acid		Q=Glutamine, R=Arginine, S=Serine,
		l		residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	Ì	l	l		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	ļ	/-possible nucleotide deletion, \-possible
	<u> </u>	 -		sequence	ļ	nucleotide insertion
		1	}			HPPPEEDQGEERPRLWVMPNHQVLLGPEEDQ DHIYHPQ*GSRGHHCPRPVPRPRLLGLGPSLP
	}	1	ł			CPS
861	2211	A	7161	1220	1003	NYVCTIAF*EKKMGF*LSLSCLVLLFVLFLDCI
601	2211	1 ^	/101	1220	1003	LTTTTRIMFHCTYLFASVCLSLLNTLLSPNCL
						KSAMILO .
862	2212	A	7211	665	847	LKYYHITMGIYKTGKKVIL*KSSMSNRFSVIF
802	2212	^	/211	003	047	
863	2213	A	7212	924	1273	YKNIQKLSFSNYVYHQNYVFSSDWSYDF HGSSCALGDLAPG*LPSGPVLSSPAVRL*RKP
603	2213	^	1212	724	1273	
		1	1	!	J	LVWDSPSCLPATGPT*GLVLVLGGPDCT*WA
		l	•		ļ	RGQHEHKRMRAP*SCRVTVNLAKKKKKTDQ
864	2214	A	7214	845	1619	CIKPNYQSPPKECDYNILANSVA SDKGGKKADRKNHLRHAFPLLPHRVRERLH
004	2214	1 ~	1214	043	1013	DPKVPVDADHVQGQDPGRAAHDIHGEDVTE
		1			ļ	KVSKDPLAPDEVGDTDEGHDRHGHREVGQR
		1				HGHDQEEVAYEERACEGGKFATVEVTDKPV
	1	ł	1	ł	ł	DEALREAMPKVAKYAGGTNDKGIGMGMTV
		1				PISFAVFPNEDGSLQKKLKVWFRIPNQFQSDP
		ļ		}	ļ	PAPSDKSVKIEEREGITVYSMQFGGYAKEAD
				1		YVAQATRLRAALEGTATYRGDIYFCTGYDPP
		1		}		MKPYGRRNEIWLLKT
865	2215	A	7246	559	682	RRLGAVAHAYTSSTLGGRGGWIT*GQELQTS
005		, .	12.10	1 337	002	LANMAKPRLY
866	2216	A	7257	641	1310	TCTYKYLMGWIRGRRSRHSWEMSEFHNYNL
,]		1	1	DLKKSDFSTRWQKQRCPVVKSKCRENASPFF
						FCCFIAVAMGIRFIIMVAIWSAVFLNSLFNQEV
		J]	j	QIPLTESYCGPCPKNWICYKNNCYQFFDESKN
:		١.		j		WYESQASCMSQNASLLKVYSKEDQDLLKLV
						KSYHWMGLVHIPTNGSWQWEDGSILSPNLLT
	١.					IIEMQKGDCALYASSFKGYIENCSTPNTYICM
						QRTV
867	2217	A	7288	151	396	SIKIIEAFGSNGPDFWFFRYWSP*LFRQQVVFI
	Ì					MPFFQTLWLMNANRFCSIFTTTNVANNCWW
		<u> </u>				TPYHCWLSVVVCRCESHGI
868	2218	Α	7298	3	272	PDTVIGGRGSGGKEFGRWVLW*VFE*RLGTP
	[1	KGSCPAGGSRMVSESD*EGRGC*ASYPCAC*
				L		AGS*WR*GSRPAGRGTPPRSLSHARPP
869	2219	Α	7332	1223	332 .	PRRDAEDRDESCLNPAFPIGLLHPNSVNSMAR
	}	1	Į.			FLTLCTWLLLLGPGLLATVRAECSQDCATCS
	1	Į.		1	ł	YRLVRPADINFLACVMECEGKLPSLKIWETC
	1	1			ļ	KELLQLSKPELPQDGTSTLRENSKPEESHLLA
		1	1	1	1	KRYGGFMKRYGGFMKKMDELYPMEPEEEA
		1	1	1	J	NGSEILAKRYGGFMKKDAEEDDSLANSSDLL
	1		i	1		KELLETGDNRERSHHQDGSDNEEEVSKRYGG
]] .	Ì	FMRGLKRSPQLKEKAKELQKRYGGFMRRVG
	[1	!	1	1	PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE
052			L	 		SYSKEVPEMEKRYGGFMRF
870	2220	Α	7382	216	1018	EIHQRLTERTQFLDESRKNPNS*QANLLRGGG
}			1	1		AGQGRGREGAESGGSRGEGPGSDGRLPATGD
				1		FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT
	l	1		1		TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL
	1	1	1		ĺ	KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF
		1	1	1		VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD
	ł	ł	1	{		FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF
	1					GFIASFMFLLDFITMLYEKRQESQLRKPENTT
871	2221	 	7402	 	202	RAEALTEPLNA
0/1	2221	A	7403	3	393	SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN
	· .	Ì		1	ļ	DPGNMSFVKETVDKLLTGFRCFREREAAPRR
	1	1		L	L	ALRGAALPGESEAGDPESLRSSVNADWIQYS

ero m	CEO ID	Mot	Teen	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	"0" -	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	Ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ľ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ł	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						DLWEAEVSTPRCEAGFCQECFRTPGNQEKDG
·		L				PFIC
872	2222	Α	7413	1061	359	FVDIVSVVEFPHCPEARFPAQHGQDSKRLTLC
						PGGS*PQATLHLDRMRVSASPTKEIQVKKYK
]	ļ	j			CGLIKPCPANYFAFKICSGAANVVGPTMCFED RMIMSPVKNNVGRGLNIALVNGTTGAVLGQ
	Ì					KAFDMYSGDVMHLVKFLKEIPGGALVLVAS
						YDDPGTKMNDESRKLFSDLGSSYAKQLGFRD
						SWVFIGAKDLRGKSPFEQFLKEQPQTQNKYE
	ł	Ì		ĺ	ľ	GWPELLEMEGCMPPKPF
873	2223	A	7429	2242	2394	ILKCAGHGGSCL*SQHFGRLRWEDRLRLGVQ
0/3	222	<u> </u>	1425	2272	2334	DHPGQHCETPSLLKIERKLF
874	2224	A	7468	146	894	PCTSCVLWATLHLPASTRKAPQAECGMISITE
0/4	2224	^h	/400	140	654	WQKIGVGITGFGIFFILFGTLLYFDSVLLAFGN
						LLFLTGLSLIIGLRKTFWFFFQRHKLKGTSFLL
	1	İ				GGVVIVLLRWPLLGMFLETYGFFSLFKGFFPV
	1	ł	ł	}	ł	AFGFLGNVCNIPFLGALFRRLQGTSSMV*KTE
		1	İ			MSSLNLDHWLKGAKREEWEPPPQSPALTHSP
	ļ	1		1		TYPGPPQVQKERNGAEQLTSNPQVDSRGCQE
						AEMQTPRRLGWGWYHTLTLYLWEEK
875	2225	Α	7498	91	251	GEKPVPTWLQDEAGQWLLGFVAQPWGWPG
		1		1		SERHEP*HGGVLFRLGPSAPPGKL
876	2226	A	7544	403	587	YSCLCFLFKHITSFKNSVHIWLGTVVHAYNPN
		-				ILGGQGGWIA*GQEFKTSLGNTVRPCLYK
877	2227	A	7566	2	940	GCAPDTRFFVPEPGGRGAAPWVALVARGGC
		1	1	}		TFKDKVLVAARRNASAVVLYNEERYGNITLP
		1	İ			MSHAGTGNIVVIMISYPKGREILELVQKGIPV
		i i			Ì	TMTIGVGTRHVQEFISGQSVVFVAIAFITMMII
						SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI
	1			1		GQLLLHTVKHGEKGIDVDAENCAVCIENFKV
						KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL
						DVIKALGYWGEPGDVQEMPAPESPPGRDPAA NLSLALPDDDGSDESSPPSASPAESEPQCDPSF
		1]		KGDAGENTALLEAGRSDSRHGGPIS
070	2220	 	7586	315	1222	ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR
878	2228	Α	/386	313	1232	RGRMQAACWYVLFLLQPTVYLVTCANLTNG
						GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS
	}	1			}	OTFRGKENDTDLDLRYDTPEPYSEODLWDW
						LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW
l		1			1	GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF
						RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID
		i .				AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK
		1			1.	TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD
		1				YKLVQKVCPDYNYHSDTPYFPSG
879	2229	A	7605	479	391	TESWKLKWWSPTCLDQLNGSAPGNVFIHG
880	2230	A	7612	93	659	DAAVAMTAQGGLVANRGRRFKWAIELSGPG
		1				GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS
						VOETDRILVEKRCWDIALGPLKQIPMNLFIMY
1	1	1				MAGNTISIFPTMMVCMMAWRPIQALMAISAT
	1	1		1		FKMLESSSQKFLQGLVYLIGNLMGLALAVYK
[1			}	CQSMGLLPTHASDWLAFIEPPERMEFSGGGL
ļ		1			· ·	LL
881	2231	A	7615	291	1452	SPOKTMRSHTITMTTTSVSSWPYSSHRMRFIT
[**	[NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT
]		1 .			1 .	SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL
1		1		1	1	LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL
]				CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI
		1		1		QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL
						TLKKGGHNSTMCFHYRDKHNAKGEAIFNFIL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VVMFWLIFLLIILSYIKIGKNLLRISKRRSKFPN SGKYATTARNSFIVLIIFTICFVPYHAFRFTYISS QLNVSSCYWKEIVHKTNEIMLVLSSFNSCLDP VMYFLMSSNIRKIMCQLLFRRFQGEPSRSEST SEFKPGYSLHDTSVAVKIQSSSKST
882	2232	A	7617	67	379	RQMALLKANKDLISAGLKEFSVLLNQQVFND PLVSEEDMVTVVEDWMNFYINYYRQQVTGE PQERDKALQELRQELNTLANPFLAKYRDFLK SHELPSHPPPSS
883	2233	A	7622	400	215	KVKTCRYNPKYSAANDTGFVDIPSREKDLAK
884	2234	A	7638	2640	2861	AVATVGPISVAVGASHVFFQFYKKGKHLSS APVLILQMVKLSIVLTPQFLSHDQGQLTKELQ QHVKSVTCPCEYLRKVSECRQMGPGALEQFP
885	2235	A	7642	201	455	GLSCHTSHSG PSRGKMELEAMSRYTSPVNPAVFPHLTVVLL AIGMFFTAWFFVYEVTSTKYTRDIYKELLISL VASLFMGFGVLFLLLWVGIYV
886	2236	A	7692	61 _	569	APENPFSRQHFNSETKVKLSLKTGTWLGNHA HLGEHFSTHHELGLSGKVVGFLVKNILEVIRN GGMETRHPGKVSSWFHRWDSRAEQHNHAE HHEDVPQGDEDSKVSEAQQEFPDVVTCAGLP GLLPKALRVLLFQLKVQHRPGIHQQRPEQQD VSDHRYGRSVRQNRK
887	2237	A	7693	85	315	NPGCCLPVAMRTSYLLLFTLCLLLSEMASGG NFLTGLGHRSDHYNCVSSGGQCLYSACPIFTK IQGTCYRGKAKCCK
888	2238	A	7702	242	1298	APSHRRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC
889	2239	A	7707	185	2911	CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA ARNNSPPTVGAFGHTRCSAFPLEQEADLIEAA

000 10	Lanam	154	T			
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		ļ		peptide	1	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1 1		ì	1			EPGGPHSSRNGLCHPLNHSRTLAGKRPKAPR
		1				GEEAHLPPVSDLTVEFDKLNLQNIGRSVSKTP
			i I		i	DESTKTKDQILTSRINAVERDLLEPSPADQLG
1		l	1		}	NGHRRTESEMSARIAKMSLSPSSPRHEDQLEV
				,		TREPARRLFLFGEEPSKLDQDVLAALECADV
						DPHQFPAVHRWKSAVLCYSPSDRQSWPSPAV
1 .		1	'			KGRFKSQLPDLSGPHSYSPGRNSVAGSNPAKP
L						GLGSPGRYSPVHGSQLRRMARLAELAAL
890	2240	Α	7711	360	269	RHMPVIPALWEAEVGGLLEPRSSRSAWATE
891	2241	Α	7721	61	1175	KLPWEPSFLIKMQIIRHSEQTLKTALISKNPVL
		Ì				VSQYEKLDAGEQRLMNEAFQPASDLFGPITL
		i			į	HSPSDWITSHPEAPQDFEQFFSDPYRKTPSPN
		1]			KRSIYIQSIGSLGNTRIISEEYIKWLTGYCKAYF
		l]			YGLRVKLLEPVPVSVTRCSFRVNENTHNLQIH
		ľ	}			AGDILKFLKKKKPEDAFCVVGITMIDLYPRDS
		1				WNFVFGQASLTDGVGIFSFARYGSDFYSMHY
			. ,:			KGKVKKLKKTSSSDYSIFDNYYIPEITSVLLLR
		ļ		<i>2</i> 1		SCKTLTHEIGHIFGLRHCQWLACLMQGSNHL
.						EEADRRPLNLCPICLHKLQCAVGFSIVERYKA
		ŀ				LVRWIDDESSDTPGATPEHSHEDNGNLPKPV
1 1		Ì	1			EAFKEWKEWIIKCLAVLQK
892	2242	A	7723	2	1650	
	2272	1.7	112	2	1000	SAPTAPARPCRAERGSGGGMLALLAASVALA
1 1		ļ				VAAGAQDSPAPGSRFVCTALPPEAVHAGCPL
						PAMPMQGGAQSPEEELRAAVLQLRETVVQQ
1						KETLASARAIRELTGKLARCEGLAGGKARGA
1						GATGKDTMGDLPRDPGHVVEQLSRSLQTLK
				ı		DRLESLEPLPAMPMQGGAQSPEELRAAVLQ
						LRETVVQQKETLASARAIRELTGKLARCEGL
j						AGGKARGAGATGKDTMGDLPRDPGHVVEQ
1						LSRSLQTLKDRLESLEHQLRANVSNAGLPGD
1						FREVLQQRLGELERQLLRKGAELEDEKSLLH
						NETSAHRQKTESTLNALLQRVTELERGNSAF
			[[KSPNAFKVSLPLRTNYLYGKIKKTLPELYAFT
						ICLWLRSSASPGMGTPFSYAVPGQANEIVLIE
						WGNNPIELLINDKVAQLPLFVSDGKWHHICV
					' ·	TWTTRDGMWEAFQDGKKLGTGENLAPWHPI
						KPGGVLILGQEQDTVGGRFDATQAFVGELSQ
				ļ		FNIWDRVLRAQEIVNIANCSTNMPGNIIPWVD
 						NNVDVFGGASKWPVETCEERLLDL
893	2243	Α	7729	3554	2419	LTAGTAMNYPLTLEMDLENLEDLFWELDRL
				[ſ	DNYNDTSLVENHLCPATEGPLMASFKAVFVP
				- 1		VAYSLIFLLGVIGNVLVLVILERHRQTRSSTET
<u> </u>				ļ		FLFHLAVADLLLVFILPFAVAEGSVGWVLGTF
	ĺ			ĺ	ļ	LCKTVIALHKVNFYCSSLLLACIAVDRYLAIV
1				j		HAVHAYRHRRLLSIHITCGTIWLVGFLLALPEI
				j	ļ	LFAKVSQGHHNNSLPRCTFSQENQAETHAWF
1	ľ		1	1		TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR
			}	ł		QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV
					l	IFLDTLARLKAVDNTCKLNGSLPVAITMCEFL
			1	ŀ		GLAHCCLNPMLYTFAGVKFRSDLSRLLTKLG
				ļ		CTGPASLCQLFPSWRRSSLSESENATSLTTF
894	2244	Ā	7738	670	287	FVTRAGRWGAGARVRGGAGGMASGAARWL
	'		50	3.0	20,	VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR
						SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI
]			VIDEL AND ALL DESIDE MEDDADA
	1					VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD
895	2245	A	7753	119	270	D APVALICATION DEVICE LIBERTORY
6/3	2243	n.	כנוו	117	278	APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR
896	2246	Ā	7754	1	270	LWLSLFLHAGKEAPHCPRTRPL
1070	2240		1154	<u> </u>	372	SPAWWNSQQRVVSPFLALLTLEPTFHHLLPIM

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QVSTAALAVLLCTMALCNQVLSAPLAADTPT ACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFL
897	2247	A	7761	1725	445	TKRGRQVCADPSEEWVQKYVSDLELSA RPRRRGTHHFSCVLGSFRVSAMFPRVSTFLPL RPLSRHPLSSGSPETSAAAIMLLTVRHGTVRY RSSALLARTKNNIQRYFGTNSVICSKKDKQSV RTEETSKETSESQDSEKENTKKDLLGIIKGMK VELSTVNVRTTKPPKRRPLKSLEATLGRLRRA TEYAPKKRIEPLSPELVAAASAVADSLPFDKQ TTKSELLSQLQQHEEESRAQRDAKRPKISFSNI ISDMKVARSATARVRSRPELRIQFDEGYDNYP GQEKTDDLKKRKNIFTGKRLNIFDMMAVTKE APETDTSPSLWDVEFAKQLATVNEQPLQNGF EELIQWTKEGKLWEFPINNEAGFDDDGSEFH EHIFLEKHLESFPKQGPIRHFMELVTCGLSKNP YLSVKQKVEHIEWFRNYFNEKKDILKESNIQF KLRPWKFLFRNN
898	2248	A	7775 •	85	496	SCQTTQPPAQSCSTGTMRIMLLFTAILAFSLA QSFGAVCKEPQEEVVPGGGRSKRDPDLYQLL QRLFKSHSSLEGLLKALSQASTDPKESTSPEK RDMHDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKYPPRAE
899	2249	Α	7785	179	703	PFHLGASSNTFRLQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEFMLMNARLQLERQLIM QSEMRERQMAMQIAWSREFLKYFGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQYDL GYGTLLERMKGEAEDILETEKSKLQLPRGMIT FESIEKARKEQSRFFIDK
900	2250	A	7789	1465	300	VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN
901	2251	A	7796	2	807	VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR
902	2252	A	7802	2	721	TAARRRQKGTAARRLQKGTAARRRQKGTAA RRRQKGTAARRPQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRRQKGLAIASRGCPCASR AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG AAACVYFGIAASLVKMGRLEGWEVFAKPKV

CECID	CEO ID	Mot	LCEO	Predicted	Deadisted and	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ		Predicted end	
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	иепсе	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ł				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	004201100	/=possible nucleotide deletion, \=possible
ļ	ļ	l)		1	nucleotide insertion
000	0052		7007	sequence		
903	2253	Α	7807	¹	584	PWLPWSDGRAARSSRKCPRSRFPVQVGKMA
ľ			1			VSTVFSTSSLMLALSRHSLLSPLLSVTSFRRFY
ļ	ļ	1	l .		. ·	RGDSPTDSQKDMIEIPLPPWQERTDESIETKR
						ARLLYESRKRGMLENCILLSLFAKEHLQHMT
			ì	ļ	1	EKQLNLYDRLINEPSNDWDIYYWATEAKPAP
		1				EIFENEVMALLRDFAKNKNKEQRLRAPDLEY
l		1	ł	i	ł	LFEKPR
904	2254	A	7813	40	821	GAGRALGHLETGAGDVAAALPARKFPRSLLG
904	2234	Α	/613	1 40	021	
ł						AGARLTGWTMNVFRILGDLSHLLAMILLLGK
ŀ		1	· .			IWRSKCCKGISGKSQILFALVFTTRYLDLFTNF
1	ļ		ļ]	}	ISIYNTVMKVVFLLCAYVTVYMIYGKFRKTF
ľ	1				i	DSENDTFRLEFLLVPVIGLSFLENYSFTLLEIL
		1				WTFSIYLESVAILPQLFMISKTGEAETITTHYL
}	-	1	1			FFLGLYRALYLANWIRRYQTENFYDQIAVVS
1	l	1	ł	ł	ĺ	GVVQTIFYCDFFYLYVTKGRSWDDSNADTGL
1.	1		1			RSYSSI
905	2255	Α	7817	1399	881	LSNKDVLSPQLKDENSKLRRKLNEVQSFSEA
903	2233	A	/61/	1399	001	
1						QTEMVRTLERKLEAKMIKEESDYHDLESVVQ
ļ		}	1	}	j	QVEQNLELMTKRAVKAENHVVKLKQEISLL
ł			1			QAQVSNFQRENEALRCGQGASLTVVKQNAD
1	1	1	1			VALQNLRVVMNSAQASIEQLVSGAETLNLVA
			1			EILKSIDRISEVKDEEEDS
906	2256	A	7822	3	1462	DSPRNRFEILGRPTRTPTRPGPRPAMEDLDAL
700		**	1.022] ~	1.02	LSDLETTTSHMPRSGAPKERPAEPLTPPPSYG
ł			1			HQPQTGSGESSGASGDKDHLYSTVCKPRSPK
ļ		1	1			
İ	1	1	i	,		PAAPAAPPFSSSSGVLGTGLCELDRLLQELNA
						TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK
1	•			1		RPSLPSSPSPGLPKASATSATLELDRLMASLSD
İ		1				FRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG
		1	i	İ		KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC
ł		l	1	ļ	}	NKPIAGQVVTALGRAWHPEHFVCGGCSTAL
[GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI
			ļ			RHKMVTALGTHWHPEHFCCVSCGEPFGDEG
	1	1	1			FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN
Į.	}	1	1	l		YISALSALWHPDCFVCRECFAPFSGGSFFEHE
l		į.				
1						GRPLCENHFHARRGSLCATCGLPVTGRCVSA
i	i		ł		:	LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY
	<u></u>	<u> </u>	<u> </u>			CQPCFLKLFG
907	2257	Α	7828	1792	1671	FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH
]						YCKSQAWG
908	2258	Α.	7842	110	1172	KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG
1	1	1		1	1	SLOPPPSGLKOSSHLSLSSSWDFRHAPTHPET
l		l	1	1	1	YTCPKMIEMEQAEAQLAELDLLASMFPGENE
1				1		LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI
į		l				
		1	1	1		NMNLDVSDEKMAMFSLACILPFKYPAVLPEI
1				[TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV
J	J	1]	Ī	}	CILNATEWVREHASGYVSRDTSSSPTTGSTVQ
		1	ł	Ī	1	SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL
I		1	Ī	l		SLSGFSMPGKPGVVCVEGPQSACEEFWARLR
1	1	i	1			KLNWKRILIRHREDIPFDGTNDETERQRKFSIF
1		ļ	1	}	l	EEKVFSVNGARGNHMDFGQLYQFLNTKGCG
1			1		ĺ	DVFQMFLWV
000	2250	 	7070	2067	2022	! — · · · · · · · · · · · · · · · · · ·
909	2259	A	7870	3067	2923	EGICVYTFIYVHMYTRTCMHTYPYMYMNSV
		L				LISSEILLIPSKYLFESK
910	2260	A	7884	212	4874	GALTWSHPLLAVCPQGVWLGSTPSGSPALLP
ĺ	1	ĺ			1	PSHRVNAEPGCVVTNACASGPCPPHANCRDL
i					1	WQTFSCTCQPGYYGPGCVDACLLNPCQNQG
1				Ī		SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD
}]]	1)	1	QQCPRGWWGSPTCGPCNCDVHKGFDPNCNK
	1	l .	I .	L .		The state of the s

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Method	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TNGQCHCKEFHYRPRGSDSCLPCDCYPVGST SRSCAPHSGQCPCRPGALGRQCNSCDSPFAEV TASGCRVLYDACPKSLRSGVWWPQTKFGVL ATVPCPRGALGLRGAGAAVRLCDEAQGWLE PDLFNCTSPAFRELSLLLDGLEINKTALDTME AKKLAQRLREVTGHTDHYFSQDVRVTARLL AHLLAFESHQQGFGLTATQDAHFNENLLWA GSALLAPETGDLWAALGQRAPGGSPGSAGLV RHLEEYAATLARNMELTYLNPMGLVTPNIML SIDRMEHPSSPRGARRYPRYHSNLFRGQDAW DPHTHVLLPSQSPRPSPSEVLPTSSSIENSTTSS VVPPAPPEPEPGISIIILLVYRTLGGLLPAQFQ AERGARLPQNPVMNSPVVSVAVFHGRNFLR GILESPISLEFRLLQTANRSKAICVQWDPPGLA EQHGVWTARDCELVHRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAAILLSLRSLKSNVRGIHANVAAA LGVAELLFLLGIHRTHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR FYHALGWGPAVLLGLAVGLDPEGYGNDF CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLVSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGPGAYNNTALFEESGLIRTLGASTVSS VSSARSGRTQDQDSGRGSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDSDLSLEEERSLSIPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGDDTSKDAAN NNQPDPALTSGDETSLGRAQRRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEFKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLFSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSRLASFNSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDAPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSRCSPRAFTHQWRTWLQCSRANAPEH LEELGRWGSAPRTHQWRTWLQCSRANAPEH LEELGRWGSAPRTHQWRTWLQCSRANAPEH LEELGRWGSAPRTHQWRTWLQCSRANAPH LEELGRWGSAPRTHQWRTWLQCSRANAPH LEELGRWGSAPRTHQWRTWLQCSRANAPH LEELGRWGSAPRTHQWRTWLQCSRANAPH LEELGRWGSAPRTHQWRTWLQCSRANAPH LEELGRWGSAPRTHQWRTWLQCSRANAPH LEELGRWGSAPRTHQ
911	2261	A	7890	21	806	
						YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD YFIYSQVTFRGMTSECSEIRQAGRPNKPDSITV VITKVTDSYPEPTQLLMGTKSVCEVGSNWFQ PIYLGAMFSLQEGDKLMVNVSDISLVDYTKE DKTFFGAFLL
912	2262	A	7891	1263	111	ACGIRHEGALPGLTATPEAMLRFLPDLAFSFL LILALGQAVQFQEYVFLQFLGLDKAPSPQKFQ PVPYILKKIFQDREAAATTGVSRDLCYVKELG VRGNVLRFLPDQGFFLYPKKISQASSCLQKLL YFNLSAIKEREQLTLAQLGLDLGPNSYYNLGP ELELALFLVQEPHVWGQTTPKPGKMFVLRSV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysinc, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PWPQGAVHFNLLDVAKDWNDNPRKNFGLFL EILVKEDRDSGVNFQPEDTCARLRCSLHASLL VVTLNPDQCHPSRKRRAAIPVPKLSCKNLCH RHQLFINFRDLGWHKWIIAPKGFMANYCHGE CPFSLTISLNSSNYAFMQALMHAVDPEIPQAV CIPTKLSPISMLYQDNNDNVILRHYEDMVVD
913	2263	A	7892	15	849	ECGCG ASRLPRGPGCGADMRPLLGLLLVFAGCTFAL YLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAE LRELSEVLREYRKEHQAYVFLLFCGAYLYKQ GFAIPGSSFLNVLAGALFGPWLGLLLCCVLTS VGATCCYLLSSIFGKQLVVSYFPDKVALLQR KVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPI LNIPIVQFFFSVLIGLIPYNFICVQTGSILSTLTS LDALFSWDTVFKLLAIAMVALIPGTLIKKFSQ KHLQLNETSTANHIHSRKDT
914	2264	Α	7893	815	959	KSGWVWWLTPLIPAL WEAQTEGSLRPEVKN RLSNITRPFFSKKKKILV
915	2265	A	7909	3	641	HASGPGGLLRRRGSGANMPVARSWVCRKT YVTPRRPFEKSRLDQELKLIGEYGLRNKREV WRVKFTLAKIRKAARELLTLDEKDPRRLFEG NALLRRLVRIGVLDEGKMKLDYILGLKIEDFL ERRLQTQVFKLGLAKSIHHAHVLIQQCHIRVR EQVVNILFFTVRLDSQKHIDFSLCFPIGVANPS HVKRKNASKGQGGAGARDDEEEE
916	2266	A	7914	3	967	VAHTQWHTCQRLSQLTHRSILKYLLIDTHAC QVLILKHTHASLSLPSCQECFPSSIPSASHMVS HPHPPPSPRWGQTPEGLPAASPCGPGPRSCFS SILPTGDSWGMLACLCTVLWHLPAVPALNRT GDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLN YLGPPFNEPDFNPPRLGAETLPRATVDLEVW RSLNDKLRLTQNYEAYSHLLCYLRGLNRQAA TAELRRSLAHFCTSLQGLLGSIAGVMAALGY PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWL LKELQTWLWRSAKDFNRLKKKMQPPAAAVT LHLGAHGF
917	2267	A	7921	2	1166	RPRRGQGLVQEVQTENVTVAEGGVAEITCRL HQYDGSIVVIQNPARQTLFFNGTRALKDERFQ LEEFSPRRVRIRLSDARLEDEGGYFCQLYTED THHQIATLTVLVAPENPVVEVREQAVEGGEV ELSCLVPRSRPAATLRWYRDRKELKGVSSSQ ENGKVWSVASTVRFRVDRKDDGGIIICEAQN QALPSGHSKQTQYVLDVQYSPTARIHASQAV VREGDTLVLTCAVTGNPRPNQIRWNRGNESL PERAEAVGETLTLPGLVSADNGTYTCEASNK HGHARALYVLVVYGESRLRPTEGGGGAPDP GAVVEAQTSVPYAIVGGILALLVFLIICVLVG MVWCSVRQKGSYLTHEASGLDEQGEAREAF LNGSDGHKRKEEFFI
918	2268	A	7938	3	2653	RRRLPPASPPSSSVSSSLSPSAVVMACRWSTK ESPRWRSALLLLFLAGVYGNGALAEHSENVH ISGVSTACGETPEQIRAPSGIITSPGWPSEYPAK INCSWFIRANPGEITTISFQDFDIQGSRRCNLD WLTIETYKNIESYRACGSTIPPPYISSQDHIWIR FHSDDNISRKGFRLAYFSGKSEEPNCACDQFR CGNGKCIPEAWKCNNMDECGDRSDEEICAKE ANPPTAAAFQPCAYNQFQCLSRFTKVYTCLP ESLKCDGNIDCLDLGDEIDCDVPTCGQWLKY FYGTFNSPNYPDFYPPGSNCTWLIDTGDHRK

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ing to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				•	·	VILRFTDFKLDGTGYGDYVKIYDGLEENPHK LLRVLTAFDSHAPLTVVSSSGQIRVHFCADKV NAARGFNATYQVDGFCLPWEIPCGGNWGCY TEQQRCDGYWHCPNGRDETNCTMCQKEEFP CSRNGVCYPRSDRCNYQNHCPNGSDEKNCFF CQPGNFHCKNNRCVFESWVCDSQDDCGDGS DEENCPVIVPTRVITAAVIGSLICGLLLVIALG CTCKLYSLRMFERRSFETQLSRVEAELLRREA PPSYGQLIAQGLIPPVEDFPVCSPNQASVLENL RLAVRSQLGFTSVRLPMAGRSSNIWNRIFNFA RSRHSGSLALVSADGDEVYPSQSTSREPERNH THRSLFSVESDDTDTENERRDMAGASGGVAA PLPQKVPPTTAVEATVGACASSSTQSTRGGH ADNGRDVTSVEPPSVSPARHQLTSALSRMTQ GLRWVRFTLGRSSSLSQNQSPLRQLDNGVSG REDDDDVEMLIPISDGSSDFDVNDCSRPLLDL ASDQGQGLRQPYNATNPGVRPSNRDGPCERC GIVHTAQIPDTCLEVTLKNETSDDEALLLC
919	2269	A	7951	1674	1839	VVRVTCCPPARSTTERTNAYDEEDCVEMVAS
920	2270	Α	7953	47	572	GGWNDVACHTTMYFMCEFDKKNM GGRASWPEQAKEPRREGHTDKQQTEDVLAA GLRCLPHLPAICARRMSPAFRAMDVEPRAKG VLLEPFVHQVGGHSCVLRFNETTLCKPLVPRE HQFYETLPAEMRKFTPQYKGKSQLLEGLPHW RGDVRDRGHGRPWQPSLEPSLPPTLCFPSLSS FSSSWPSAQHLTPSVFNPW
921	2271	A	7957	612	812	RSGRTVVTGIGYSKALQSSNRNTKSLLQNEF MMVYSFRALSFKESTWATFQHGGEATKSRSL SSTQ
922	2272	A	7967	1443	1660	ENITEK WKEIWMCRGNKKSCCWTFIKDRHLT VSCCKSKSGETLLICIFCSNLVGFFFFGIRGFSN WELVKPN
923	2273	A	7981	1	3023	GSAPRAATAMARARPPPPPSPPPGLLPLLPLL LLPLLLPAGCRALEETLMDTKWVTSELAWT SHPESGWEEVSGYDEAMNPIRTYQVCNVRES SQNNWLRTGFIWRRDVQRVYVELKFTVRDC NSIPNIPGSCKETFNLFYYEADSDVASASSPFW MENPYVKVDTIAPDESFSRLDAGRVNTKVRS FGPLSKAGFYLAFDQQGACMSLISVRAFYKK CASTTAGFALFPETLTGAEPTSLVIAPGTCIPN AVEVSVPLKLYCNGDGEWMVPVGACTCATG HEPAAKESQCRPCPPGSYKAKQGEGPCLPCPP NSRTTSPAASICTCHNNFYRADSDSADSACTT VPSPPRGVISNVNETSLILEWSEPRDLGVRDD LLYNVICKKCHGAGGASACSRCDDNVEFVPR QLGLSEPRVHTSHLLAHTRYTFEVQAVNGVS GKSPLPPRYAAVNITTNQAAPSEVPTLRLHSS SGSSLTLSWAPPERPNGVILDYEMKYFEKSEG IASTVTSQMNSVQLDGLRPDARYVVQVRART VAGYGQYSRPAEFETTSERGSGAQQLQEQLP LIVGSATAGLVFVVAVVVIAIVCLRKQRHGS DSEYTEKLQQYIAPGMKVYIDPFTYEDPNEA VREFAKEIDVSCVKIEEVIGAGEFGEVCRGRL KQPGRREVFVAIKTLKVGYTERQRRDFLSEA SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME NCALDSFLRLNDGQFTVIQLVGMLRGIAAGM KYLSEMNYVHRDLAARNILVNSNLVCKVSDF GLSRFLEDDPSDPTYTSSLGGKIPIRWTAPEAI AYRKFTSASDVWSYGIVMWEVMSYGERPY

NO. of NO. of load in loading nucleicide cotide of sequence peptide sequence unce unce unce unce unce unce unce	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nuclectide sequence USSN 09496 (ocation corresponding sequence) 914 (ocation corresponding to the properties of the period residue of peptide sequence) 914 (ocation can be form to the peptide sequence) 914 (ocation can be form to the peptide sequence) 915 (ocation can be form to the peptide sequence) 915 (ocation can be form) 91							D=Aspartic Acid, E=Glutamic Acid,
Sequence				in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
1	eotide	1	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
amino acid residue of peptide residue of peptide residue of peptide sequence T=Threonine, V=Vaine, W=Typtophan, Y=Typtophan, Y=Sep codon, V=possible nucleotide deletion, V=possible nucleotide deletion, V=possible nucleotide insertion VDMSNQDVINAPEDVRIPPMDCPTALHQ LINLOWVERDRIN.RPKS:GIVINIT.DALIR.NRA SLKVIASA; SGMSGPLIDRIVPDYTIFTTY GO	seq-	uence		09/496	correspondi	to last amino	
Peptide Pept	uence		ļ	914		acid residue	
peptide	1		{	i		of peptide	
			İ			sequence	
WIMSINGDVINAVEGDYRILERINA LMLDCWYRDRINERFSQIVITLUKLIRNAA SLKVIASAQSGMSQPLLDRIVPDYTFTTYGD WIDAIKMGRYKESFYSQIVATIOLA/AQMTA EDILRIGVTLAGHQKKILSSIQDMRLQMNQT LPVQV 924 2274 A 7985 1 563 FEPRITKATAMYLEHYLDSEIN_PCELORNF EDILRIGVTLAGHQKKILSSIQDMRLQMNQT LPVQV 925 2275 A 7994 447 589 LPCSFCAQCMSSFELVWLQQSHFHNPRWNSR EDIPFKKKHKGG SDFESSGGRILKKGRQKKRKGRGRGRTS EDIPFKKKHKGG EDIPFKKKHKGG PCCCCTILAIMLQCHSFYRKDVQVEHPKS LNFKYSQLENISADMALKRCLLSISDLDFW IWDAQPVGMQTLQNIKKPPHGCVENVQAFG RDTQPLIPLGGRYYTTIRQ RDTQPLIPLGGRYYTTIRQ PCCCTILAIMLQCHSFYRKDVQVEHPKS LNFKYSQLENISADMALKRCLLSISDLDFW IWDAQPVGMQTLQNIKKPPHGCVENVQAFG RDTQPLIPLGGRYYTTIRQ RDTQPLIPLGGRYYTTIRQ PCCCTILAIMLQCHSFYRKACHASCAC RDTQPLIPLGGRYYTTIRQ PCCCTILAIMLQCHSFYRKACHASCAC RDTQPLIPLGGRYYTTIRQ PCCCTILAIMLQCHSFYRKACHASCAC RDTQPLIPLGGRYYTTIRQ PCCCTILAIMLQCHSFYRKACHASCAC RDTQPLIPLGGRYYTTIRQ PCCCTILAIMLQCHSFYRKACHASCAC RDTQPLIPLGGRYYTTIRQ PCCCTILAIMLQCHSFYRKACHASCAC RDTQPLIPLGGRYYTTIRQ PCCCTILAIMLQCHSFYRKACHASCAC RDTQPLIPLGGRYYTTIRQ PCCCTILAIMLQCHSFYRACHASCAC RDTQPLIPLGGRYYTTIRQ PCCCTILAIMLQCHSFYRACHASCAC RDTQPLIPLGGRYYTTIRQ PCCCTILAIMLQCHSFYRACHASCAC RDTQPLIPLGGPTTYFWAPTAWACHASCAC RDTQPLIPLGGPTTYFWAPTAWACHASCAC PCCCTILAIMLGCHSFYRACHASCAC RAGLARTYFHLIDKVELMPERLAPLYNIPLGAC PCCCTILAIMLGCHSFYRACHASCAC RDTQPLIPLGGPTTYFWAPTAWACHACHASCAC PCCCTILAIMLGCHSFTACHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMCCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMCCHASCAC PCCCTILAIMLGCHASCAC PCCCTILAIMCCHASCAC PCCCTILAIMCCHASCAC PCCCTILAIMCCHASCAC PCCCTILAIMCCHASCAC PCCCTILAIMCCHASCAC PCCCTILAIMCCHASCAC PCCCTILAIMCCHASCAC PCCCTILAIMCCHASCAC PCCCTILAIMCCHASCAC PCCCTILAIMCCHASCAC PCCCTILAIMC				l			
LMLDCWVRDRNLRPKFSQIVNTLDKLIRNAA SLKVIJSAQSGMSQPLLDRIVPDYTTFTTYGD WILDAIRMGRYKESFVSAGFASFDLVAQMTA SLKVIJSAQSGMSQPLLDRIVPDYTTFTYGO WILDAIRMGRYKESFVSAGFASFDLVAQMTA EDLLRIGHTAGHOKKILSSIQDMRLQMNQT LPVQV PFPFTKKATAMYLEHYLDSENLPCELORNF QLMRELDQRTEDKKAEDILAAEYISTVKTLS PDQRVERLQKIQNAYSKCKEYSDDKVQLAM QTYEMVDKHIRRLDADLAAFF2ALIXOKMEG SDFESSGGRGLKKGRGQKEKKGRGRGRATIS EDTPFKKKHKGG PFCFKCKKHKGG PFCFKCCMSFERVWLQQSHFHNDRWNSE PFCFKCAQCMSSFERVWLQQSHFHNDRWNSE PFCFKCAQCMSSFERVWLQQSHFHNDRWNSE PFCFKCAQCMSSFERVWLQQSHFHNDRWNSE PFCFKCAQCMSSFERVWLQQSHFHNDRWNSE PFCFKCAQCMSSFERVWLQQSHFHNDRWNSE PFCFKCAQCMSFERVWLQQSHFHNDRWNSE PFCFKCAQCMSFERVWQQSHFHNDRWNSE PFCFKCAQCMSFERVWQQSHFHNDRWNSE PFCFKCAQCMSFERVWQQSHFHNDRWNSE PFCFKCAQCMSFERVWQQSHFHNDRWNSE PFCFKCAQCMSFERVWQQSHFHNDRWNSE PFCFKCAQCMSFERVWQQSHFHNDRWNSE PFCFKCAQCMSFERVWQQSHFHNDRWNSE PFCFKCAQCMSFERVWQQSHFHNDRWNSE PFCFKCAQCMSFERVWQQSHFHNDRWNSE PFCFKCAQCMSFERVWQQSHFHNDRWNSE PFCFKCAQCMSFERVWQQSHFHNDRWNSE PFCFKCAQCMSFERVWQQAACMAG PFCFKCAQCMSFERVWQQAACMAG PFCFKCAQCMSFERVWQQAACMAG PFCFKCAQCMSFERVWQQAACMAG PFCFKCAQCMSFAAKUACMACACMAG PFCFKCAQCMSFAAKUACMACACMAG PFCFKCAQCMSFAAKUACMACACMAG PFCFKCAQCMSFAAKUACMAG PFCFKCAQCMSFAAKUACMAG PFCFKCAQCMSFAAKUACMAG PFCFKCAQCMSFAAKUACMAG PFCFKCAQCMSFAAKUACMAG PFCFKCAQCMSFAAKUACMAG PFCFKCAQCMSFAAKUACMAG PFCFKCAACMAG PF		<u>L</u>	<u> </u>	<u></u>	sequence		
	1			}		1	
			1	1			LMLDCWVRDRNLRPKFSQIVNTLDKLIRNAA
LPVQV		ļ	1	1			
924	ļ		ļ	}	1-	l	,
Polymerion Pol	004	2274	 	7005		602	
POQRVERI_OKIQNAYSKCKEYSDDK_VOLAM	924	2274	A	/985	1	503	
925 2275 A 7994 447 589				İ			
SDFESSGGRGLKKGRGQKEKRGSRGRGRRTS							
	{		[ĺ			
925 2275 A 7994 447 589 LPCSFCAQCMSSFERVWLQQSHFHNPRWNSR 926 2276 A 7996 925 582 GPCKVCCTILAIMLQCHSFYRKDVQVEHPKS LNPKYSQIENPLSADMALKRCLISISDLDFW INDAQPVIGNIQTIQNLKKIPPGGFWSQAFQI RDTQPILPLGGRYYTITRQ 927 2277 A 7998 2 353 RIGPENGSSESPHISLFVKABLTAKQATMKLSV CLLLVTLALCCYQANAEFCPALVSELLDFFI SEPLFKLSLAKFDAFPEAVAKLGVKRCTDQ MSIQKRSLLAKFDAFPEAVAKLGVKRCTDQ 928 2278 A 8004 130 588 LAPLRCQPGTRTQPRSHPANDFSAMISAAG GRAGLATYHRILDKVEL MLPEKLRPLYNIPA GPRTYFFWAPIMK WGLVCAGLADMARPAEK LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA ANGELRQSPRHTEYAGPRISASASKELLM KLRKKTGYSFVNCKKALBTCGGDLKQAEIWL HKKEAQKEGWSKAAKLLOPKRINGELKAKAHK KLRKKTGYSFVNCKKALBTCGGDLKQAEIWL HKKEAQKEGWSKAAKLOPKRISAGSSESLGLPA GPDREGSLKDQLALAIGKKTKEGLIGLLQE GNITVLVEVNCETDFVSRNLKFQLLVQQVAL GTMHICQTILKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKKTKEGLIGLLDE GNITVLVEVNCETDFVSRNLKFQLLVQWAL GTMHICQTILKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKSGNETKEGLIGLLQE GNITVLVEVNCETDFVSRNLKFQUHLGWAPL SVGSLDDEPGGRAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVFREGGEAAETE GLUBURGWSCSLLVASIGNGSSFLYGYN LSVVNAPTPYIKAFYNESWERHGRPDPDIT. TILLWSVTVSRAGLQVTILKMIGKVLGRK HILLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSSIGHAGRAF GRAVKAFQTFLGKADVSQCVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGINAIWFYTNSIFGKAGPFAKIPYVTLSTGG IETLAAVFSGLVIEHLGRFRLLIGGFGIMGLFF GTILTITLILQHAPWVPYLSVGLIALASFCSG PGGFPITLGEFFTQSQCRFAAFIRAGTVNWLSN FAVGLLEFFTQSQCRFAAFIRAGTVNWLSN <td></td> <td></td> <td></td> <td>1</td> <td></td> <td></td> <td></td>				1			
SPIRCYCQHWPHCVHC	025	2275	<u> </u>	700/	447	580	
926 2276 A 7996 925 582 GPCKVCTILAMI_QCHSFYRKDVQVEHPKS_LNPKYSQIENFLSADMALKRKCLLSISDLDFW_IDAQPVGIMQTI_QNI_KKIPNPGCFWSQAFQI_RDTOPILP_IGGRYTITIRQ 927 2277 A 7998 2 353 RIQRPLINSRSPNHSLIFVKAELTAKQATMKLSV_CLI_LVTLALCCYQANAEFCPALVSELLDFFFI_SEPL_FKLSLAKFDAPPSAVAKLIGVKRCTDQ 928 2278 A 8004 130 588 LAPLRCQPGTRTQPRSHPAANDPSAAMSAAG 929 2279 A 8007 2 1016 EFARRYFTAAREMSLLERLRYLVARTGSYP 4 4 8007 2 1016 EFARRYFTAAREMSLLERLRYLVARTGSYP 4 4 8007 2 1016 EFARRYFTAAREMSLLERLRYLVARTGSYP 4 4 8007 2 1016 EFARRYFTAAREMSLLRSLRYLVARTGSYP 4 4 8007 2 1016 EFARRYFTAAREMSLLRSLRYLVARTGSYP 4 4 4 8007 2 1016 EFARRYFTAAREMSLLRSLRYLVARTGSYP 4 4 4 4 4 4 4 4 4 4 4 4 4	923	2213	1	1774	***/	369	
LNPKYSQIENFISADMALKRKCLISISDLDFW WDAQPYOGMOTICQHKKPNPGCFWSQAFQI RDTQPILPLGGRYYTTIRQ 927 2277 A	926	2276	Δ	7996	925	582	
1 1 1 1 1 1 1 1 1 1	320	2270	^	1770	723	302	
927 2277 A 7998 2 353 RIQRPLNSRSPNHSLFVKAELTAKQATMKLSY CLLLVTLALCCYQANAEFCPALVSELLDFFFI SEPLFKLSLAKFDAPPEAVAAKLGVKRCTDQ MSLQKKSLIAEVLKKCSV 928 2278 A 8004 130 588 LAPLRCQPGTRTQPRSHPAANDPSAAMSAAG ARGLRATYHRLDKVELMLPEKLRPLYNHPA GPRTVFFWAPIMK WGLVCAGLADMARPAEK LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA VNFFVQAAGASQLFMRYNYQELKAKAKK LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA VNFFVQAAGASQLFMRYNYQELKAKAKK LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA VNFFVQAAGASQLFMRYNYQELKAKAKK LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA VNFFVQAAGASQLFMRYNYQELKAKAKK LGRARTKEGSFPNCKKALETGGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETDYSRRLKFQLVQQVAL GTMMHCQTIKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWK KLPSGFFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLQGHVVGMAPL SVGSLDDEPGGBAETKMLSQPYLIDPSITLGQ YVQPQGVSVVDFVFFECGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGFWTEPSAGSIRPMARKQNRNSKEL GLVPLTDDTSHAGPPCFGRALLECDHLRSCV PGGRRKGWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFPSWFRRHGRPDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTILLANNGFAISAALLMACSLQAGAFEMLIV GRFMGIDGGVALLYMYLSSFFERGSLG QVTAFFICIGVFTQLLGLPELLGKESTWPYLF GVTVVPAVVQLLSLPFLPDSPRYLLEKHNEA RAVKAFQTFLGKAGIPPAKPYVTLSTGG IETLAAVFSGLVIEHLGGRPLMGFFYTLSFFINGST GIFTLAAVFSGLVIEHLGRRPLLIGGGGLMGLFF GTLTITILTLQDHAPWVPYLISVGILAIASFCSG PGGIFFILGRFPQSQRPAAFIIAGTVNWLSN FAVGLLFFFIQKSLDTYCFLVFATTCITGATYL , YFVLFETKNRTVAGQAFSKRNKAYPPEEK DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKKL NNNARRKAAPRIECSHIRHAWDHAKSVRQNL			į		ļ		1 1 - 1 - 1 - 1 - 1
927			l		Ì		
2278 A 8004 130 588 LAPIRCQPGTRTQPSHPAANDPSAAMSAAG ARGLRATYHRLLDKVELKKCSV 928 2278 A 8004 130 588 LAPIRCQPGTRTQPSHPAANDPSAAMSAAG ARGLRATYHRLLDKVELMLPKLRPLYNHPA GPRTVFFWAPIMKWGLVCAGLADMARPAEK LSTAQSAVLMATGFIWRSYSLVIPKNWSLFA VNFFVQAAGASQLFIWRYNQELKAKAHK 929 2279 A 8007 2 1016 EFARRRVFIAAREMSLLRSLRVFLVARTGSYP AGSLLRQSPQPRHTTYAGPRLSASASSKELLM KLRRKTGYSFVNCKALETICGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGULQE GNTTVLVEVNCETDFVSRIKGLLVQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRIGQHVVGMAPL SVGSLDDEFGGEAETKMLSQYYLLDESITLGQ YQPQGVSVVDFYPEGGEGEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRIGRPDPDTIL TILWSVVTVSIFAIGGLVGTLIVKMIGKVLGRK HTILANNGFAISAALLMACSLQAGAFFMLIV GRFINGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLEFLLGKESTWPYLF GGTVVTPAVVQLLSLPFLPDSRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVVELLRAPVVPVSIVGLAILASFCSG PGGIPFLTGEFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAVIL TYVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NNNARRKAAPRIEGSHIRHAWDHAKSVRQNL NNNARRKAAPRIEGSHIRHAWDHAKSVRQNL NNNARRKAAPRIEGSHIRHAWDHAKSVRQNL NNNARRKAAPRIEGSHIRHAWDHAKSVRQNL NNNARRKAAPRIEGSHIRHAWDHAKSVRQNL NNNARRKAAPRIEGSHIRHAWDHAKSVRQNL NNNARRKAAPRIEGSHIRHAWDHAKSVRQNL	927	2277	A	7998	2	353	
	121		**	1770	-	1 323	
MSI_QKRSLIAEVI_VKILKKCSV					•		
928 2278 A 8004 130 588 LAPLRCOPGTRTOPRSHPANDPSAAMSAAG ARGLRATYHRILDKVELMLPEKLRPLYNHPA GPRIVFFWAPHMWGUCAGLADMARPAEK LSTAQSAVLMATGFIWSRYSLVIPKNWSLFA VNIFFVGAAGASQLFRIWRYNQELKAKAHK 929 2279 A 8007 2 1016 FEARRWIJARAREMSLLRSLRVFLVARTGSYP AGSLLRQSPOPRHTFYAGPRLSASASSKELLM KLRRKTYSFVNCKKALETCGGDLKAEIWL HKEAQKEGWSKAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA AVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFCGEGEEAAETTE 930 2280 A 8008 3 1679 NSRVWGPWTFPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAAGPPGFGRALLECDHLRSGV PGGRRRCDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYKAFYNESWERRHGRPIDDPDTL TLLWSVTVSFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGYTGQLLGLPELLGKESTWPYLF GVVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVTVIMACYQL CGLNAIWFYTNSIFGKAGIPAKIFYTLSTGG IETLAAVFSGLVIELGRPAKIFYTLSTGA AGAVVSAMPKAKGKTRQKFGYSVNRKRL NRARRKAAPHECSHITHAWDHAKSVQNL				i		}	
ARGIRATYHRLLÖKVELMI.PEKLRPLYNHPA GPRTVFFWAPIMK WGLVCAGLADMARPAEK LSTAQSAVLMATOFIWSRYSL VIIPKNWSLFA VNFFVGAAGASQLFRIWRYNQELKAKAHK 929 2279 A 8007 2 1016 EFARRRVFIAAREMSLLRSLRVFLVARTGSYP AGSLLRQSPQPRHTFYAGPRLSASASSKELLM KLRKKTGYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTILDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGBAETKMLSQPYLLDPSTILGQ YVQPQGVSVVDFVRFECGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRFMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGKK HTILANNGFAISAALLMACSLQAGAFEMLIV GRFMGIDGGVALSVLPMYLSEISFKEIRGSLG QVTAIFICIGVFTGQLLGLEPLICKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQVEEVLAESRVQRS IRLVSVLELILRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETTLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITILTLQDHAPWVPYLSIVGILAILASSCSG PGGIPFLTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLPFFIQKSLDTYCFLVFATICITGAIYL YFYLPETKNRTYAEISQAFSKRIKAYPPEEKI DSAVTTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARKAAAPRIECSSHIRIAWDHAKSVRQNL	928	2278	A	8004	130	588	LAPLRCQPGTRTQPRSHPAANDPSAAMSAAG
LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA			1				
929 2279 A 8007 2 1016 EFARRVFIAAREMSLLRSLRVFLVARTGSYP AGSLLRQSPQRHTIFYAGFRLSASASKELLM KLRKTGYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQRKFKEGLIGLLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLAAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEFGGEATKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGGEEAAETE 930 2280 A 8008 3 1679 NSRVWGFWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAFTYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTILLANNGFAISAALLMACSLQAGAFEMLIV GRFMGIDGGVALSVLPMYLSEISFKEIRGSLG QVTAIFICIGVFTGQLLGLEGLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAYGLLFFPGKSLDTVCFLVFATICTGAIFL, YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL			İ	į	:		GPRTVFFWAPIMKWGLVCAGLADMARPAEK
929 2279 A 8007 2 1016 EFARRRVFIAARÉMSLLRSLRVFLVARTGSYP AGSILRQSPOPRHITYAGPRI.SASASSKELLM KLRKTGYSFVNCKKALETCGGDLKQAEIUM KLRRKTGYSFVNCKKALETCGGDLKQAEIUM KLRRKTGYSFVNCKKALETCGGDLKQAEIUM KLRRKTGYSFVNCKKALETCGGDLKQAEIUM KLRRKTGYSFVNCKKALETCGGDLKQAEIUM GTMMHCQTILKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRALGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYILDPSITLGQ YVQPQGVSVVDFVRFECGEGEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYKAFYNESWERRHGRPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTILANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GYIVVPAVVQLLSLPFLPDSPRYLILEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGHNAIWFYTNSIFGKAGJPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITTLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLPFFQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKNGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			1				LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA
AGSLLRQSPQPRHTIFYAGPRLSASASSKELLM KLRRKTGYSFYNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTILANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAMWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITITLIQDHAPWVPYLSIVGILAIIASFCSG PGGPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLPFFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKNGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL]	}			J		VNFFVGAAGASQLFRIWRYNQELKAKAHK
KLRKTGYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLEPLLGKESTWPYLF GGTVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAYKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGPPAKIPYYTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	929	2279	Α	8007	2	1016	
HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETDFYSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVFFECGEGEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNFGAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTTFLGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRPLLIGGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFLTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFYLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			1]	
GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVFFECGEGEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYTKAFYNESWERHGRPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGKK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAFFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFTLTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL							
GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSSFYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVEHLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			į.				
GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGREPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAITASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		1	[·	1			
KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWYPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		1	ł		Ì		
ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLS1PFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVVRQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVEHLGRPPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		1	İ				
SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFLTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL							
930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKNIGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		ł	ł	1	1	}	
930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERHGRPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGYFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL							
GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	020	2000	 	0000	3	1670	
PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELRAPYVRWQVVTVITMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	A20	2280	A	0008	3	10/9	
LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	1		1	1	1	1	
TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL							
HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			1		1		
GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	1	1	1		1	1	
QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	1	j	J]	1	
GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGGRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL					1	1	1
RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			İ		1	-	
IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			I				
CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	1	1			l	1	
IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDT YCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	1	1	1		1	1	
GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDT YCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		1	1		1		
PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDT YCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			1	1	1		
FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		1			1		
YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	1		ł	1	1	1	
931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		1.	1		Į.		
931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		1	1		1		
NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	931	2281	A	8009	861	300	
]				1		
, I I I I I I I I I I I I I I I I I I I	1		1	1	1	1	AEMGLAVDPNRAVPLRKRKVKAMEVDIEER

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		İ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
,			İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l		peptide		/=possible nucleotide deletion, \=possible
,		l		sequence		nucleotide insertion
						PKELVRKPYVLNDLEAEASLPEKKGNTLSRD
						LIDYVRYMVENHGEDYKAMARDEKNYYOD
				ĺ		TPKQIRSKINVYKRFYPAEWQDFLDSLQKRK
						MEVE
932	2282	Α	8011	412	1	SNLCLGNSWRWRWAKSRHHCIPTVTLSKRSG
		1	ļ.	1		DIRGSHFSSPQRQRSQRVPGKETARVLRAGK
				1		QGRGQIPIPCPWPPPPPPPPPPGSPGPGCRQFHQ
	ĺ			ļ		SLEAKARHPASVREMRGKVKMRRALRRAPA
						STRASSROPNPK
933	2283	Α	8012	147	1077	PPVPPASRSDMAQNLKDLAGRLPAGPRGMGT
		1			1	ALKLLLGAGAVAYGVRESVFTVEGGHRAIFF
		1	1	1		NRIGGVQQDTILAEGLHFRIPWFQYPIIYDIRA
]	ł		}	j	RPRKISSPTGSKDLQMVNISLRVLSRPNAQEL
		1	1			PSMYQRLGLDYEERVLPSIVNEVLKSVVAKF
		1	1		ļ	NASQLITQRAQVSLLIRRELTERAKDFSLILDD
		1		1		VAITELSFSREYTAAVEAKQVAQQEAQRAQF
			i	1		LVEKAKOEOROKIVOAEGEAEAAKMLGEAL
	ļ	}	j	}	J	SKNPGYIKLRKIRAAQNISKTIATSQNRIYLTA
	•					DNLVLNLQDESFTRGSDSLIKGKK
934	2284	A	8023	255	982	SQFSLSQVLVDSAEEGSLAAAAELAAQKREQ
754	2204	^	8023	255.	702	RLRKFRELHLMRNEARKLNHQEVVEEDKRL
	1	Į	1	1		KLPANWEAKKARLEWELKEEEKKKECAARG
				İ		EDYEKVKLLEISAEDAERWERKKKRKNPDLG
	}					FSDYAAAQLRQYHRLTKQIKPDMETYERLRE
İ	1			1		KHGEEFFPTSNSLLHGTHVPSTEEIDRMVIDLE
		ĺ			Í	
		1		1		KQIEKRDKYSRRRPYNDDADIDYINERNAKF NKKAERFYGKYTAEIKQNLERGTAV
935	2285	A	8027	59	310	LVSSTVNLLTEKAPWNSLAWTVTSYVFLKFL
733	2205	A	0027	1 39	310	
					1	QGGGTGSTGMRDSALTLLGIGPSHRHSLSIRL
936	2286	A	8032	1	639	SQHSSPAPMYSQTFHILVLG
930	2200	Α	0032	1	039	SGRECNMAKTYDYLFKLLLIGDSGVGKTCVL
	ı		l		i	FRESEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ
					ļ	IWDTAGQERFRTITTAYYRGAMGIMLVYDIT
			1 '			NEKSFONIRNWIRNIEEHASADVEKMILGNKC
		Į	1			DVNDKRQVSKERGEKLALDYGIKFMETSAK
	[.	ĺ	İ			ANINVENAFFTLARDIKAKMDKKLEGNSPQG
027	2207	 	10020	202	211	SNQGVKITPDQQKRSSFFRCVLL
937	2287	A	8039	393	311	EETIHSENSYILEKYIPISANLTLTIA
938	2288	A	8052	675	-1334	LHPAATSTAWLHVPPGLSMALSWVLTVLSLL
			1	!		PLLEAQIPLCANLVPVPITNATLDRITGKWFYI
ļ	1	[1	ſ	ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF
	1		1	1		LREYQTRQDQCIYNTTYLNVQRENGTISRYV
						GGQEHFAHLLILRDTKTYMLAFDVNDEKNW
	1	I	1		[GLSVYADKPETTKEQLGEFYEALDCLRIPKSD
	ļ			L		VVYTDWKKDKCEPLEKQHEKERKQEEGES
939	2289	Α	8055	12	1039	SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF
	1	1	1			AEQLKWSAELARLGESIMDGKQGGMDGSKP
	1	ł	1	l	1	AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA
			1			IHGHQLSLRNLISQGWAVNIITADHVSPLHEA
	1		ļ	1		CLGGHLSCVKILLKHGAQVNGVTADWHTPL
	1			1		FNACVSGSWDCVNLLLQHGASVQPESDLASP
•			1	1	1	IHEAARRGHVECVNSLIAYGGNIDHKISHLGT
•	ł	[
					İ	PLYLACENOORACVKKLLESGADVNOGKGO
						PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTOAKN
						DSPLHAVARTASEELACLLMDFGADTQAKN
-						DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL
-						DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH
940	2290	A	8058	2	1203	DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL

	T			 	T & T	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ł			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	1	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
]	ľ	peptide	1	/=possible nucleotide deletion, \=possible
		l		sequence	1	nucleotide insertion
	 	 	┼			VVDTVMCPNMPNKSVLLYTLSFIYIFIFVIGMI
	1		}]		ANSVVVWVNIQAKTTGYDTHCYILNLAIADL
	i				Į	WVVLTIPVWVVSLVQHNQWPMGELTCKVTH
	Į.		İ		l	LIFSINLFGSIFFLTCMSVDRYLSITYFTNTPSS
	ŀ					
]	l]		RKKMVRRVVCILVWLLAFCVSLPDTYYLKT
						VTSASNNETYCRSFYPEHSIKEWLIGMELVSV
			1 .	1		VLGFAVPFSIIAVFYFLLARAISASSDQEKHSS
			ł			RKIIFSYVVVFLVCWLPYHVAVLLDIFSILHYI
	j -	· .	•			PFTCRLEHALFTALHVTQCLSLVHCCVNPVL
	1			İ	ł	YSFINRNYRYELMKAFIFKYSAKTGLTKLIDA
		į				SRVSETEYSALEQSTK
941	2291	Α	8059	73	432	DMAGLMTIVTSLLFLGVCAHHIIPTGSVVLPS
	j	ļ		ļ	}	PCCMFFVSKRIPENRVVSYOLSSRSTCLKAGV
						IFTTKKGQQFCGDPKQEWVQRYMKNLDAKQ
		Ì				KKASPRARAVAVKGPVQRYPGNQTTC
942	2292	A	8067	278	1262	GGIGEIKORPSCLGRCLDPSLSVLMNISLGLGS
742	2272	A	0007	270	1202	VFSAVISQKPSRDICQRGTSLTIQCQVDSQVT
	1					MMFWYRQQPGQSLTLIATANQGSEATYESGF
	1		İ			
	1	1		Ì		VIDKFPISRPNLTFSTLTVSNMSPEDSSIYLCSA
	ł					GRQGTYEQYFGPGTRLTVTEDLKNVFPPEVA
				ł		VFEPSEAEISHTQKATLVCLATGFYPDHVELS
						WWVNGKEVHSGVSTDPQPLKEQPALNDSRY
	(i		}		CLSSRLRVSATFWQNPRNHFRCQVQFYGLSE
						NDEWTQDRAKPVTQIVSAEAWGRADCGFTS
	1	1				ESYQQGVLSATILYEILLGKATLYAVLVSALV
						LMAMVKRKDSRG
943	2293	Α	8070	1	879	MVKVVPATRGNLPRSQLTGTHQHCQPREPKI
						TASERLRRRPRATARLRAHAAPPEPPLAVFAP
	1					PSDRKELLALPVACDPVIASVMSWVQAASLI
	1					QGPGDKGDVFDEEADESLLAQREWQSNMQR
			ĺ			RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA
						EVILNYGRLRGTLSALLSWCHLHNNNSTLINK
						INNLLDAVGOCEEYVLKHLKSITPPSHVVDLL
						DSIEDMDLCHVVPAEKKIDEAKDERLCENNA
	l	Ì	1			1
	1					EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK
044	0004	<u> </u>	0050	<u> </u>		PKPHMDFGTDSQF
944	2294	A	8073	1	797	ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK
	1	ĺ	1			MAATSGTDEPVSGELVSVAHALSLPAESYGN
		ŀ				DPDIEMAWAMRAMQHAEVYYKLISSVDPQF
						LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK
		l				SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD
	1	l				CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA
	Í	1				VYISVQDKEGEKGVNNGGEKRADSGEEENT
	1	l				KNGGEKGADSGEEKEEGINREDKTDKGGEK
		1				GKEADKEINKSGEKAM
945	2295	Α	8074	2	505	GAATLLRSASSAARKAAEAEQVWLHLHRYL
· · ·		١	55.4] ~		SADRRVLGLREWGRPASERECSLCQRLKREL
	1	l]			
	İ	!	l .			NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT
1		1	1	[GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW
			1			GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER
		<u> </u>				ADLIAYLKKATNE
946	2296	Α	8081	42	590	EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF
		ĺ	1		[VAIFAVPLILGQEYEDEERLGEDEYYQVVYY
	1	ļ				YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK
	İ					DITEAIETTISLETARADHPKPVTVKPVTTEPQ
						SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC
	ĺ	[1			KKVGRRLLMTLWMGVWOEEIGR
947	2297	A	8084	322	549	GGGSSPRELAGAAGLTVTSQAVAARRQQPSF
	'			1		SRARAPAHSLRAALSLASSARSWGAVSRDRG
		L		ı		

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide Medical N=Control N
948	2298	В	8093	3905	846	PCPPAIMYQSSNKC MEPGEVKDRILENISLSVKKLQSYFAACEDEI PAIRNHDKVLQRLCEHLDHALLYGLQDLSSG YWVLVVHFTRREAIKQIEVLQHVATNLGRSR AWLYLALNENSLESYLRLFQENLGLLHKYYV KNALVCSHDHLTLFLTLVSGLEFIRFELDLDA PYLDLAPYMPDYYKPQYLLDFEDRLPSSVHG SDSLSLNSFNSVTSTNLEWDDSAIAPSSEDYD FGDVFPAVPSVPSTDWEDGDLTDTVSGPRST ASDLTSSKASTRSPTQRQNPFNEEPAETVSSS DTTPVHTTSQEKEEAQALDPPDACTELEVIRV TKKKKIGKKKKSRSDEEASPLHPACSQKKCA KQGDGDSRNGSPSLGRDSPDTMLASPQEEGE GPSSTTESSERSEPGLLIPEMKDTSMERLGQPL SKVIDQLNGQLDPSTWCSRAEPPDQSFRTGSP GDAPERPPLCDFSEGLSAPMDFYRFTVESPST VTSGGGHHDPAGLGQPLHVPSSPEAAGQEEE GGGGEGQTPRPLEDTTREAQELEAQLSLVRE GPVSEPEPGTQEVLCQLKRDQPSPCLSSAEDS GVDEQQGSPSEMVHSSEFRVDNNHLLLLMIH VFRENEEQLFKMIRMSTGHMEGNLQLLYVLL TDCYVYLLRKGATEKPYLVEEAVSYNELDY VSVGLDQQTVKLVCTNRRKQFLLDTADVAL AEFFLASLKSAMIKGCREPPYPSILTDATMEK LALAKFVAQESKCEASAVTVRFYGLVHWED PTDESLGPTPCHCSPPEGTITKEGMLHYKAGT SYLGKEHWKTCFVVLSNGILYQYPDRTDVIP LLSVNMGGEQCGGCRRANTTDRPHAFQVILS DPPCLELSAESEAEMAEWMQHLCQAVSKGVI PQGVAPSPCIPCCLVLTDDRLFTCHEDCQTSF FRSLGTAKLGDISAVSTEPGKEYCVLEFSQDS QQLLPPWVIYLSCTSELDRLLSALNSGWKTIY QVDLPHTAIQEASNKKKFEDALSLIHSAWQR SDSLCRGRASRDPWC*
949	2299	A	8095	9	2374	ARRADTVLLESPSMLQGLLPVSLLLSVAVSAI KELPGVKKYEVVYPIRLHPLHKREAKEPEQQ EQFETELKYKMTINGKLAVLYLKKNKNLLAP GYTETYYNSTGKEITTSPQIMDDCYYQGHILN EKVSDASISTCRGLRGYFSQGDQRYFIEPLSPI HRDGQEHALFKYNPDEKNYDSTCGMDGVL WAHDLQQNIALPATKLVKLKDRKVQEHEKY IEYYLVLDNGEFKRYNENQDEIRKRVFEMAN YVNMLYKKLNTHVALVGMEIWTDKDKIKIT PNASFTLENFSKWRGSVLSRRKRHDIAQLITA TELAGTTVGLAFMSTMCSPYSVGVVQDHSD NLLRVAGTMAHEMGHNFGMFHDDYSCKCPS TICVMDKALSFYIPTDFSSCSRLSYDKFFEDKL SNCLFNAPLPTDIISTPICGNQLVEMGEDCDC GTSEECTNICCDAKTCKIKATFQCALGECCEK CQFKKAGMVCRPAKDECDLPEMCNGKSGNC PDDRFQVNGFPCHHGKGHCLMGTCPTLQEQ CTELWGPGTEVADKSCYNRNEGGSKYGYCR RVDDTLIPCKANDTMCGKLFCQGGSDNLPW KGRIVTFLTCKTFDPEDTSQEIGMVANGTKCG DNKVCINAECVDIEKAYKSTNCSSKCKGHAV CDHELQCQCEEGWIPPDCDDSSVVFHFSIVVG VLFPMAVIFVVVAMVIRHQSSREKQKKDQRP LSTTGTRPHKQKRKPQMVKAVQPQEMSQMK PHVYDLPVEGNEPPASFHKDTNALPPTVFKD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
950	2300	A	8100	1	1251	MGLLLMILASAVLGSFLTLLAQFFLLYRRQPE PPADEAARAGEGFRYIKPVPGLLLREYLYGG GRDEEPSGAAPEGGATPTAAPETPAPPTRETC YFLNATILFLFRELRDTALTRRWVTKKIKVEF EELLQTKTAGRLLEGLSLRDVFLGETVPFIKTI RLVRPVVPSATGEPDGPEGEALPAACPEELAF EAEVEYNGGFHLAIDVDLVFGKSAYLFVKLS RVVGRLRLVFTRVPFTHWFFSFVEDPLIDFEV RSQFEGRPMPQLTSIIVNQLKKIIKRKHTLPNY KIRFKPFFPYQTLQGFEEDEEHIHIQQWALTE GRLKVTLLECSRLLIFGSYDREANVHCTLELS SSVWEEKQRSSIKTGTISLTAVFMGWHRVSE AFPGLWYKLLVDLPFWGLEDGGPLLTVPLRQ CPG
951	2301	A	8108	1612	839	EVALFCFEMAAGMYLEHYLDSIENLPFELQR NFQLMRDLDQRTEDLKAEIDKLATEYMSSAR SLSSEEKLALLKQIQEAYGKCKEFGDDKVQL AMQTYEMVDKHIRRLDTDLARFEADLKEKQI ESSDYDSSSSKGKKKGRTQKEKKAARARSKG KNSDEEAPKTAQKKLKLVRTSPEYGMPSVTF GSVHPSDVLDMPVDPNEPTYCLCHQVSYGE MIGCDNPDCSIEWFHFACVGLTTKPRGKWFC PRCSQERKKK
952	2302	A	8112	595	291	PSVASLARRFSGRALWPPSHSVPGNRALCPRL LHGTTLPGGNQRELARQKNMKKQSDSVKGK RRDDGLSAAARKQRDSTPRDSEIMQQKQKK ANEKKEEPK
953	2303	A	8118	1	669	VCAGIRDPCSTPLAKPAAGGAENLSFGKQPG LETNILKMTTPNKTPPGADPKQLERTGTVREI GSQAVWSLSSCKPGFGVDQLRDDNLETYWQ SDGSQPHLVNIQFRRKTTVKTLCIYADYKSDE SYTPSKISVRVGNNFHNLQEIRQLELVEPSGW IHVPLTDNHKKPTRTFMIQIAVLANHQNGRD THMRQIKIYTPVEESSIGKFPRCTTIDFMMYRS IR
954	2304	A	8133	66	1015	PPLPPRSFPNLFSRPEPLPEPGRRGCNRSREPA ARAPSPPPPFEGAPGRAMVKVTFNSALAQKE AKKDEPKSGEEALIIPPDAVAVDCKDPDDVV PVGQRRAWCWCMCFGLAFMLAGVILGGAY LYKYFALQPDDVYYCGIKYIKDDVILNEPSAD APAALYQTIEENIKIFEEEEVEFISVPVPEFADS DPANIVHDFNKKLTAYLDLNLDKCYVIPLNT SIVMPPRNLLELLINIKAGTYLPQSYLIHEHMV ITDRIENIDHLGFFIYRLCHDKETYKLQRRETI KGIQKREASNCFAIRHFENKFAVETLICS
955	2305	A	8143	35	1171	VESRSAWHEGEDQIDRLDFIRNQMNLLTLDV KKKIKEVTEEVANKVSCAMTDEICRLSVLVD EFCSEFHPNPDVLKIYKSELNKHIEDGMGRNL ADRCTDEVNALVLQTQQEIIENLKPLLPAGIQ DKLHTLIPCKKFDLSYNLNYHKLCSDFQEDIV FRFSLGWSSLVHRFLGPRNAQRVLLGLSEPIF QLPRSLASTPTAPTTPATPDNASQEELMITLVT GLASVTSRTSMGIIVGGVIWKTIGWKLLSVS LTMYGALYLYERLSWTTHAKERAFKQQFVN YATEKLRMIVSSTSANCSHQVKQQIATTFARL CQQVDITQKQLEEEIARLPKEIDQLEKIQNNS KLLRNKAVQLENELENFTKQFLPSSNEES
956	2306	A	8157	1854	798	ASGSPAPSSSSAMAAACGPGAAGYCLLLGLH LFLLTAGPALGWNDPDRMLLRDVKALTLHY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DRYTTSRRLDPIPQLKCVGGTAGCDSYTPKVI QCQNKGWDGYDVQWECKTDLDIAYKFGKT VVSCEGYESSEDQYVLRGSCGLEYNLDYTEL GLQKLKESGKQHGFASFSDYYYKWSSADSC NMSGLITIVVLLGIAFVVYKLFLSDGQYSPPP YSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQ NTGHGATSGFGSAFTGQQGYENSGPGFWTGL GTGGILGYLFGSNRAATPFSDSWYYPSYPPSY PGTWNRAYSPLHGGSGSYSVCSNSDTKTRTA SGYGGTRRR
957	2307	Α	8159	1492	528	THVVMTGMCYAPHQVLSYINGVTTSKPGVSL VYSMPSRNLSLRLEGLQEKDSGPYSCSVNVQ DKQGKSRGHSIKTLELNVLVPPAPPSCRLQGV PHVGANVTLSCQSPRSKPAVQYQWDRQLPSF QTFFAPALDVIRGSLSLTNLSSSMAGVYVCKA HNEVGTAQCNVTLEVSTGPGAAVVAGAVVG TLVGLGLLAGLVLLYHRRGKALEEPANDIKE DAIAPRTLPWPKSSDTISKNGTLSSVTSARAL RPPHGPPRPGALTPTPSLSSQALPSPRLPTTDG AHPQPISPIPGGVSSSGLSRMGAVPVMVPAQS OAGSLV
958	2308	A	8161	2340	1192 ·	ELARRPKQQSSEKSRNMIRNWLTIFILFPLKLV EKCESSVSLTVPPVVKLENGSSTNVSLTLRPP LNATLVITFEITFRSKNITILELPDEVVVPPGVT NSSFQVTSQNVGQLTVYLHGNHSNQTGPRIR FLVIRSSAISIINQVIGWIYFVAWSISFYPQVIM NWRRKSVIGLSFDFVALNLTGFVAYSVFNIGL LWVPYIKEQFLLKYPNGVNPVNSNDVFFSLH AVVLTLIIIVQCCLYERGGQRVSWPAIGFLVL AWLFAFVTMIVAAVGVITWLQFLFCFSYIKL AVTLVKYFPQAYMNFYYKSTEGWSIGNVLL DFTGGSFSLLQMFLQSYNNDQWTLIFGDPTK FGLGVFSIVFDVVFFIQHFCLYRKRPGYDQLN
959	2309	A	8163	521	1345	GERAGRRGRLGVWAQPQPLLPRPVGSRRE MQPPGPPPAYAPTNGDFTFVSSADAEDLSGSI ASPDVKLNLGGDFIKESTATTFLRQRGYGWL LEVEDDDPEDNKPLLEELDIDLKDIYYKIRCV LMPMPSLGFNRQVVRDNPDFWGPLAVVLFFS MISLYGQFRVVSWIITIWIFGSLTIFLLARVLG GEVAYGQVLGVIGYSLLPLIVIAPVLLVVGSF EVVSTLIKLFGVFWAAYSAASLLVGEEFKTK KPLLIYPIFLLYIYFLSLYTGV
960	2310	A	8167	1	2921	MTCFKGQKGEQRSHAFEANKDHKAKVPSPN LYSQLNALQFTVDERSILWLNQFLLDLKQSL NQFMAVYKLNDNSKSDEHVDVRVDGLMLK FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM SVDLSHIPLKDPLLFKSASDTNLQKGISFMDY LSDKHLGKISEDESSGLVYKSGSGEIGSETSD KKDSFYTDSSSVLNYREDSNILSFDSDGNQNI LSSTLTSKGNETIESIFKAEDLLPEAASLSENL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DISKEETPPVRTLKSQSSLSGKPKERCPPNLAP
		-				LCVSYKNMKRSSSQMSLDTISLDSMILEEQLL ESDGSDSHMFLEKGNKKNSTTNYRGTAESVN AGANLQNYGETSPDAISTNSEGAQENHDDLM SVVVFKITGVNGEIDIRGEDTEICLQVNQVTP DQLGNISLRHYLCNRPVGSDQKAVIHSKSSPE ISLRFESGPGAVIHSLLAEKNGFLQCHIENFST EFLTSSLMNIQHFLEDETVATVMPMKIQVSNT KINLKDDSPRSSTVSLEPAPVTVHIDHLVVER SDDGSFHIRDSHMLNTGNDLKENVKSDSVLL TSGKYDLKKQRSVTQATQTSPGVPWPSQSAN FPEFSFDFTREQLMEENESLKQELAKAKMAL AEAHLEKDALLHHIKKMTVE
961	2311	A	8172	1442	682	TAAMSIFTPTNQIRLTNVAVVRMKRAGKRFEI ACYKNKVVGWRSGVEKDLDEVLQTHSVFVN VSKGQVAKKEDLISAFGTDDQTEICKQILTKG EVQVSDKERHTQLEQMFRDIATIVADKCVNP ETKRPYTVILIERAMKDIHYSVKTNKSTKQQA LEVIKQLKEKMKIERAHMRLRFILPVNEGKKL KEKLKPLIKVIESEDYGQQLEIVCLIDPGCFREI DELIKKETKGKGSLEVLNLKDVEEGDEKFE
962	2312	A	8175	286	587	NISNKAEVSSHPSVISHSMDSFGQPRPEDNQS VLRRMQKKYWKTKQVFIKATGKKEDEHLVA SDAFLDAKLEVFHSVQETCTELLKIIEKYQLR LNGMKS
963	2313	Ą	8181	13	2215	AEGCAERRGTEPVVELSMSWESGAGPGLGSQ GMDLVWSAWYGKCVKGKGSLPLSAHGIVV AWLSRAEWDQVTVYLFCDDHKLQRYALNRI TVWRSRSGNELPLAVASTADLIRCKLLDVTG GLGTDELRLLYGMALVRFVNLISERKTKFAK VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF
964	2314	A	8184	6	1393	EPRRNFRDDSTRPRTRGRTRGRRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS WVIDQKKPVRFYHDWNDKEIEVLNKHLFLTS KPMVYLVNLSEKDYIRKKNKWLIKIKEWVD KYDPGALVIPFSGALELKLQELSAEERQKYLE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion ANMTQSALPKIKAGFAALQLEYFFTAGPDEV RAWTIRKGTKAPQAAGKIHTDFEKGFIMAEV MKYEDFKEEGSENAVKAAGKYRQQGRNYIV EDGDIIFFKFNTPQQPKKK
						AAALGGHPLLGVSATLNSVLNSNAIKNLPPPL GGAAGHPGSAVSAAPGILYPGGNKYQTIDNY QPYPCAEDEECGTDEYCASPTRGGDAGVQIC LACRKRRKRCMRHAMCCPGNYCKNGICVSS DQNHFRGEIEETITESFGNDHSTLDGYSRRTT LSSKMYHTKGQEGSVCLRSSDCASGLCCARH FWSKICKPVLKEGQVCTKHRRKGSHGLEIFQ RCYCGEGLSCRIQKDHHQASNSSRLHTCQRH
966	2316	A	8207	416	4082	KFKLIKIMLLTLIILLPVVSKFSFVSLSAPQHW SCPEGTLAGNGNSTCVGPAPFLIFSHGNSIFRI DTEGTNYEQLVVDAGVSVIMDFHYNEKRIY WVDLERQLLQRVFLNGSRQERVCNIEKNVSG MAINWINEEVIWSNQQEGIITVTDMKGNNSHI LLSALKYPANVAVDPVERFIFWSSEVAGSLY RADLDGVGVKALLETSEKITAVSLDVLDKRL FWIQYNREGSNSLICSCDYDGGSVHISKHPTQ HNLFAMSLFGDRIFYSTWKMKTIWIANKHTG KDMVRINLHSSFVPLGELKVVHPLAQPKAED DTWEPEQKLCKLRKGNCSSTVCGQDLQSHLC MCAEGYALSRDRKYCEGNDWKYCEDVNEC AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPD GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPL SPVSWECDCFPGYDLQLDEKSCAASGPQPFL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKIYFAHTALKWIERANMDGSQ RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKIITIENISQPRGIAVHPMAK RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYVWFSDWAMP SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGHHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHLREDDHHYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ
967	2317	A	8210	3	601	SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM DVEHKMSIRILK

NO: of nucl- cotide seq- uence NO: of nucl- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence USSN 10 MO: of nucle- cotide seq- uence M=Methionine, N=Asparagine, P=Proline, Q=Glutamic, G=Glycine, H=Histidine, l=Islocucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, N=Proline, Q=Glutamic, G=Glycine, H=Histidine, l=Islocucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamic, M=Lyeucine, M=Lyeucine, M=Methionine, M=Asparagine, P=Proline, O=Clutamic, R=Asparatic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Islocucine, K=Lysine, L=Leucine, M=L	IL OP IG VR GN YP RR LD DR YF
ectide sequence uence Sequence 1914	IL OP IG VR GN YP RR LD DR YF
sequence Sequence 19/496	IL OP IG VR GN YP RR LD DR YF
uence 914 ng to first amino acid residue of peptide residue of peptide sequence Po	IL OP IG VR GN YP RR LD DR YF
amino acid residue of peptide sequence peptide sequence peptide sequence s	IL OP IG VR GN YP RR LD DR YF
residue of peptide sequence /=possible nucleotide deletion, \=	IL OP IG VR GN YP RR LD DR YF
peptide sequence /=possible nucleotide deletion, \=possible nucleotide insertion 968 2318 A 8211 2 409 ISSCPHTAYEGSMSTLSNFTQTLEDVFRRIFI	IL OP IG VR GN YP RR LD DR YF
Sequence nucleotide insertion	IL OP IG VR GN YP RR LD DR YF
968 2318 A 8211 2 409 ISSCPHTAYEGSMSTLSNFTQTLEDVFRRIFI YMDNWRQNTTAEQEALQAKVDAENFYYV YLMV,MIGMFSFIIVAILVSTVKSKREHSNI YHQYIVEDWQEKYKSQILNLEESKATIHEN AAGFKMSP 969 2319 A 8215 1 1938 GMPRSRGGRAAPGPPPPPPPPPGQAPRWSRV VPGRLLLLLPALCCLPGAARAAAAAGA RAAVAVAVARADEAEAPFAGQNWLKSYG LLPYDSRASALHSAKALQSAVSTMQQFYG VTGVLDQTTIEWMKKPRCGVPDHPHLSRR NKRYALTGQKWRQKHITYSIHNYTPKVGE TRKAIRQAFDVWQKVTPLTFEEVPYHEIKS KEADIMIFFASGFHGDSSPFDGEGGFLAHA PGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPLGDRPSTP KPNICDGNFNTVALFRGEMFVFKDRWFWF RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	IL OP IG VR GN YP RR LD DR YF
YMDNWRQNTTAEQEALQAKVDAENFYYY YLMVMIGMFSFIIVAILVSTVKSKRREHSNI YHQYIVEDWQEKYKSQILNLEESKATIHEN AAGFKMSP 969 2319 A 8215 1 1938 GMPRSRGGRAAPGPPPPPPPPPPPPPPPPPPPPPPPPPPPPP	IL OP IG VR GN YP RR LD DR YF
YLMVMIGMFSFIIVAILVSTVKSKRREHSNI YHQYIVEDWQEKYKSQILNLEESKATIHEN AAGFKMSP 969 2319 A 8215 1 1938 GMPRSRGGRAAPGPPPPPPPPPQAPRWSRV VPGRLLLLLLPALCCLPGAARAAAAAGA RAAVAVARADEAEAPFAGQNWLKSYG LLPYDSRASALHSAKALQSAVSTMQQFYG VTGVLDQTTIEWMKKPRCGVPDHPHLSRR NKRYALTGQKWRQKHITYSHINYTPKVGE TRKAIRQAFDVWQKVTPLTFEEVPYHEIKS KEADIMIFFASGFHGDSSPFDGEGGFLAHA PGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEFTRPLF PVRRIHSPSERKHERQPRPPPPLGDRPSTPG KPNICDGNFNTVALFRGEMFVFKDRWFWW RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	OP IG VR GN Y IP RR LD DR YF
YHQYIVEDWQEKYKSQILNLEESKATIHEN AAGFKMSP 969 2319 A 8215 1 1938 GMPRSRGGRAAPGPPPPPPPPGQAPRWSRV VPGRLLLLLPALCCLPGAARAAAAAAGA RAAVAVARADEAEAPFAGQNWLKSYG LLPYDSRASALHSAKALQSAVSTMQQFYG VTGVLDQTTIEWMKKPRCGVPDHPHLSRR NKRYALTGQKWRQKHITYSIHNYTPKVGE TRKAIRQAFDVWQKVTPLTFEEVPYHEIKS KEADIMIFFASGFHGDSSPFDGEGGFLAHA PGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPPLGDRPSTPC KPNICDGNFNTVALFRGEMFVFKDRWFWF RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	IG VR GN Y IP RR LD DR YF
AAGFKMSP 969 2319 A 8215 1 1938 GMPRSRGGRAAPGPPPPPPPPGQAPRWSRV VPGRLLLLLPALCCLPGAARAAAAAAGA RAAVAVAVARADEAEAPFAGQNWLKSYG LLPYDSRASALHSAKALQSAVSTMQQFYG VTGVLDQTTIEWMKKPRCGVPDHPHLSRR NKRYALTGQKWRQKHITYSHNYTPKVGE TRKAIRQAFDVWQKVTPLTFEEVPYHEIKS KEADIMIFFASGFHGDSSPFDGEGGFLAHAPGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPLGDRPSTPKPNICDGNFNTVALFRGEMFVFKDRWFWFRNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	VR GN Y IP RR LD DR YF
969 2319 A 8215 1 1938 GMPRSRGGRAAPGPPPPPPPPGQAPRWSRV VPGRLLLLLPALCCLPGAARAAAAAAGA RAAVAVAVARADEAEAPFAGQNWLKSYG LLPYDSRASALHSAKALQSAVSTMQQFYG VTGVLDQTTIEWMKKPRCGVPDHPHLSRR NKRYALTGQKWRQKHITYSIHNYTPKVGE TRKAIRQAFDVWQKVTPLTFEEVPYHEIKS KEADIMIFFASGFHGDSSPFDGEGGFLAHA PGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPPLGDRPSTPK KPNICDGNFNTVALFRGEMFVFKDRWFWF RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	GN Y IP RR LD DR YF
VPGRLLLLLPALCCLPGAARAAAAAGAR RAAVAVAVARADEAEAPFAGQNWLKSYG LLPYDSRASALHSAKALQSAVSTMQQFYG VTGVLDQTTIEWMKKPRCGVPDHPHLSRR NKRYALTGQKWRQKHITYSIHNYTPKVGE TRKAIRQAFDVWQKVTPLTFEEVPYHEIKS KEADIMIFFASGFHGDSSPFDGEGGFLAHAPGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPPLGDRPSTPK KPNICDGNFNTVALFRGEMFVFKDRWFWFRNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	GN Y IP RR LD DR YF
RAAVAVARADEAEAPFAGQNWLKSYG LLPYDSRASALHSAKALQSAVSTMQQFYG VTGVLDQTTIEWMKKPRCGVPDHPHLSRR NKRYALTGQKWRQKHITYSIHNYTPKVGE TRKAIRQAFDVWQKVTPLTFEEVPYHEIKS KEADIMIFFASGFHGDSSPFDGEGGFLAHA PGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPPLGDRPSTPG KPNICDGNFNTVALFRGEMFVFKDRWFWF RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	Y IP RR LD DR YF
LLPYDSRASALHSAKALQSAVSTMQQFYG VTGVLDQTTIEWMKKPRCGVPDHPHLSRR NKRYALTGQKWRQKHITYSIHNYTPKVGE TRKAIRQAFDVWQKVTPLTFEEVPYHEIKS KEADIMIFFASGFHGDSSPFDGEGGFLAHA PGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPPLGDRPSTP KPNICDGNFNTVALFRGEMFVFKDRWFWF RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	IP RR LD DR YF
VTGVLDQTTIEWMKKPRCGVPDHPHLSRR NKRYALTGQKWRQKHITYSIHNYTPKVGE TRKAIRQAFDVWQKVTPLTFEEVPYHEIKS KEADIMIFFASGFHGDSSPFDGEGGFLAHA' PGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPPLGDRPSTP- KPNICDGNFNTVALFRGEMFVFKDRWFWF- RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	RR LD DR YF
NKRYALTGQKWRQKHITYSIHNYTPKVGE TRKAIRQAFDVWQKVTPLTFEEVPYHEIKS KEADIMIFFASGFHGDSSPFDGEGGFLAHA' PGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPPLGDRPSTP' KPNICDGNFNTVALFRGEMFVFKDRWFWF' RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	LD DR YF
TRKAIRQAFDVWQKVTPLTFEEVPYHEIKS KEADIMIFFASGFHGDSSPFDGEGGFLAHA' PGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPPLGDRPSTP KPNICDGNFNTVALFRGEMFVFKDRWFWF RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	DR YF
KEADIMIFFASGFHGDSSPFDGEGGFLAHA PGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPPLGDRPSTPC KPNICDGNFNTVALFRGEMFVFKDRWFWF RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQC	YF
PGPGIGGDTHFDSDEPWTLGNANHDGNDL VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPPLGDRPSTPG KPNICDGNFNTVALFRGEMFVFKDRWFWF RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	1
VAVHELGHALGLEHSSDPSAIMAPFYQYM HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPPLGDRPSTPC KPNICDGNFNTVALFRGEMFVFKDRWFWF RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQC	FT. İ
HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLF PVRRIHSPSERKHERQPRPPRPPLGDRPSTPG KPNICDGNFNTVALFRGEMFVFKDRWFWF RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	
PVRRIHSPSERKHERQPRPPRPPLGDRPSTPG KPNICDGNFNTVALFRGEMFVFKDRWFWF RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	
KPNICDGNFNTVALFRGEMFVFKDRWFWF RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	
RNNRVQEGYPMQIEQFWKGLPARIDAAYE ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	
ADGRFVFFKGDKYWVFKEVTVEPGYPHSL ELGSCLPREGIDTALRWEPVGKTYFFKGER WRYSEERRATDPGYPKPITVWKGIPQAPQG	
WRYSEERRATDPGYPKPITVWKGIPQAPQG	_
WRYSEERRATDPGYPKPITVWKGIPQAPQG	Y
FISKEGYYTYFYKGRDYWKFDNQKLSVEP	GY
PRNILRDWMGCNQKEVERRKERRLPQDDV	DI I
MVTINDVPGSVNAVAVVIPCILSLCILVLVV	TI
FQFKNKTGPQPVTYYKRPVQEWV	
970 2320 A 8216 1235 2223 SRLSLQFYVSFRRTGLFTCKLIVEIFFRNYM	
DSLRTNVFVRFQPETIACACIYLAARALQIF	
TRPHWFLLFGTTEEEIQEICIETLRLYTRKK	
YELLEKEVEKRKVALQEAKLKAKGLNPDO	
ALSTLGGFSPASKPSSPREVKAEEKSPISINV	
TVKKEPEDRQQASKSPYNGVRKDSKRSRN	
SASRSRSTRSRSHTPRRHYNNRRSRSG	
SSRSRSRSRSHHNHGSPHLKAKH	
DDLKSSNRHGHKRKKSRSRSQSKSRDHSD KKHRHERGHHRDRRERSRSFERSHKSKHH	
· · · · · · · · · · · · · · · · · · ·	uu
971 2321 A 8217 3 3274 DCRLQAAMPTNFTVVPVEAHADGGGDET	A IC
971 2321 A 8217 3 3274 DCRLQAAMPTNFTVVPVEAHADGGGDET. RTEAPGTPEGPEPERPSPGDGNPRENSPFLN	
VEVEQESFFEGKNMALFEEEMDSNPMVSS	
NKLANYTNLSQGVVEHEEDEESRRREAKA	
MGTFIGVYLPCLQNILGVILFLRLTWIVGV	
VLESFLIVAMCCTCTMLTAISMSAIATNGV	
AGGSYYMISRSLGPEFGGAVGLCFYLGTTI	
GAMYILGTIEIFLTYISPGAAIFQAEAAGGE	
AMLHNMRVYGTCTLVLMALVVFVGVKY	
KLALVFLACVVLSILAIYAGVIKSAFDPPDI	
CLLGNRTLSRRSFDACVKAYGIHNNSATSA	•
WGLFCNGSQPSAACDEYFIQNNVTEIQGIP	GΑ
ASGVFLENLWSTYAHAGAFVEKKGVPSVI	
AEESRASTLPYVLTDIAASFTLLVGIYFPSV	TG
IMAGSNRSGDLKDAQKSIPTGTILAIVTTSF	
LSCIVLFGACIEGVVLRDKFGEALQGNLVI	
LAWPSPWVIVIGSFFSTCGAGLQTLTGAPR	
QAIARDGIVPFLQVFGHGKANGEPTWALL	
VLICETGILIASLDSVAPILSMFFLMCYLFV	
ACAVQTLLRTPNWRPRFKFYHWTLSFLGN	
CLALMFICSWYYALSAMLIAGCIYKYIEYR	_
AEKEWGDGIRGLSLNAARYALLRVEHGPF KNWRPQVLVMLNLDAEQAMKHPRLLSFT	.G
	G HT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LKAGKGLTIVGSVLEGTYLDKHMEAQRAEE NIRSLMSTEKTKGFCQLVVSSSLRDGMSHLIQ SAGLGGLKHNTVLMAWPASWKQEDNPFSW KNFVDTVRDTTAAHQALLVAKNVDSFPQNQ ERFGGGHIDVWWIVHDGGMLMLLPFLLRQH KVWRKCRMRIFTVAQVDDNSIQMKKDLQMF LYHLRISAEVEVVEMVENDISAFTYERTLMM EQRSQMLKQMQLSKNEQEREAQLIHDRNTAS HTAAAARTQAPPTPDKVQMTWTREKLIAEK YRSRDTSLSGFKDLFSMKPDQSNVRRMHTAV KLNGVVLNKSQDAQLVLLNMPGPPKNRQGD ENYMEFLEVLTEGLNRVLLVRGGGREVITTYS
714	<i>L34</i> 6	•	0224	/01	240	TSRRVTMKFNPFVTSDRSKNRKRHFNAPSHV RRKIMSSPLSKELRQKYNVRSMPIRKDDEVQ VVRGHYKGQQIGKVVQVYRKKYVIYIERVQ REKANGTTVHVGIHPSKVVITRLKLDKDRKKI LERKAKSRQVGKEKGKYKEELIEKMQE
973	2323	A	8237	279	4610	GCPHAGGKGRVPTGGLTGGRTWSPSAAPRSC PRPGPTPAPGAMDKLPPSMRKRLYSLPQQVG AKAWIMDEEDAEEGAGGRQDPSRRSIRLR PLPSPSPSAAAGGTESRSSALGAADSEGPARG AGKSSTNGDCRFRGSLASLGSRGGGSGGTG SGSSHGHLHDSAEERRLIAEGDASPGEDRTPP GLAAEPERPGASAQPAASPPPPQQPPQPASAS CEQPSVDTAIKVEGGAAAGDQILPEAEVRLG QAGFMQRQFGAMLQPGVNKFSLRMFGSQKA VEREQERVKSAGFWIIHPYSDFRFYWDLTML LLMVGNLIIIPVGITFFKDENTTPWIVFNVVSD TFFLIDLVLNFRTGIVVEDNTEIILDPQRIKMK YLKSWFMVDFISSIPVDYIFLIVETRIDSEVYK TARALRIVRFTKILSLLRLLRLSRLIRYIHQWE EIFHMTYDLASAVVRIVNLIGMMLLLCHWDG CLQFLVPMLQDFPDDCWVSINNMVNNSWGK QYSYALFKAMSHMLCIGYGRQAPVGMSDV WLTMLSMIVGATCYAMFIGHATALIQSLDSS RRQYQEKYKQVEQYMSFHKLPPDTRQRIHD YYEHRYQGKMFDEESILGELSEPLREEIINFNC RKLVASMPLFANADPNFVTSMLTKLRFEVFQ PGDYIIREGTIGKKMYFIQHGVVSVLTKGNKE TKLADGSYFGEICLLTRGRRTASVRADTYCR LYSLSVDNFNEVLEEYPMMRRAFETVALDRL DRIGKKNSILLHKVQHDLNSGVFNYQENEIIQ QIVQHDREMAHCAHRVQAAASATPTPPVIW TPLIQAPLQAAAATTSVAIALTHHPRLPAAIFR PPPGSGLGNLGAGQTPRHLKRLQSLIPSALGS ASPASSPSQVDTPSSSSFHIQQLAGFSAPAGLS PLLPSSSSSPPPGACGSPSAPTPSAGVAATTIA GFGHFHKALGGSLSSSDSPLLTPLQPGARSPQ AAQPSPAPPGARGGLGLPEHFLPPPPSSRSPSS SPGQLGQPPGELSLGLATGPLSTPETPPRQPEP PSLVAGASGGASPVGFTPRGGLSPPGHSPGPP RTFPSAPPRASGSHGSLLLPPASSPPPPQVPQR RGTPPLTPGRLTQDLKLISASQPALPQDGAQT LRRASPHSSGESMAAFPLFPRAGGGSGGSGSS GGLGPPGRPYGAIPGQHVTLPRKTSSGSLPPP LSLFGARATSSGGPPLTAGPQREPGARPEPVR SKLPSNL
974	2324	A	8247	279	468	EYKQWERRFLSCQNRNDLGYGKPRKGGGLL LVPVKDASRICSLTYLLGSHWNNLVVRSPVL G

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
975	2325	A	8249	62	1571	LVALKNWKPKGTNIPAPQSPVFGEAVSGVYM MTKVLGMAPVLGPRPPQEQVGPLMVKVEEK EEKGKYLPSLEMFRQRFRQFGYHDTPGPREA LSQLRVLCCEWLRPEIHTKEQILELLVLEQFLT ILPQELQAWVQEHCPESAFEAVTLLEDLEREL DEPGHQVSTPPNEQKPVWEKISSSGTAKESPS SMQPQPLETSHKYESWGPLYIQESGEEQEFAQ DPRKVRDCRLSTQHEESADEQKGSEAEGLKG DIISVIIANKPEASLERQCVNLENEKGTKPPLQ EAGSKKGRESVPTKPTPGERRYICAECGKAFS NSSNLTKHRRTHTGEKPYVCTKCGKAFGASSHLK HQRMHTEEAPYQCKDCGKAFSGKGSLIRHYR IHTGEKPYQCNECGKSFSQHAGLSSHQRLHT GEKPYKCKECGKAFNHSSNFNKHHRIHTGEK PYWCHHCGKTFCSKSNLSKHQRVHTGEGEA
976	2326	A	8257	298	7086	GNMACWPQLRLLWKNLTFRRQTCQLLLE VAWPLFIFLILISVRLSYPPYEQHECHFPNKAM PSAGTLPWVQGIICNANNPCFRYPTPGEAPGV VGNFNKSIVARLFSDARRLLLYSQKDTSMKD MRKVLRTLQQIKKSSSNLKLQDFLVDNETFS GFLYHNLSLPKSTVDKMLRADVILHKVFLQG YQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP REKLAAAERVLRSNMDILKPILRTLNSTSPFPS KELAEATKTLLHSLGTLAQELFSMRSWSDMR QEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEG GGLKIKSLNWYEDNNYKALFGGNGTEEDAE TFYDNSTTPYCNDLMKNLESSPLSRIIWKALK PLLVGKILYTPDTPATRQVMAEVNKTFQELA VFHDLEGMWEELSPKIWTFMENSGEMDLVR MLLDSRDNDHFWEQQLDGLDWTAQDIVAFL AKHPEDVQSSNGSVYTWREAFNETNQAIRTIS RFMECVNLNKLEPIATEVWLINKSMELLDER KFWAGIVFTGITPGSIELPHHVKYKIRMGIDN VERTNKIKDGYWDPGPRADPFEDMRYVWGG FAYLQDVVEQAIIRVLTGTEKKTGVYMQQMP YPCYVDDIFLRVMSRSMPLFMTLAWIYSVAV IIKGIVYEKEARLKETMRIMGLDNSILWFSWFI SSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV FLSVFAVVTILQCFLISTLFSRANLAAACGGII YFTLYLPYVLCVAWQDYVGFTLKIFASLLSP VAFGFGCEYFALFEEQGIGVQWDNLFESPVE EDGFNLTTSVSMMLFDTFLYGVMTWYIEAVF PGQYGIPRPWYFPCTKSYWFGEESDEKSHPGS NQKRISEICMEEEPTHLKLGVSIQNLVKVYRD GMKVAVDGLALNFYEGQITSFLGHNGAGKT TTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQ NLGVCPQHNVLFDMLTVEEHIWFYARLKGLS EKHVKAEMEQMALDVGLPSSKLKSKTSQLS GGMQRKLSVALAFVGGSKVVILDEPTAGVDP YSRRGIWELLLKYRQGRTIILSTHHMDEADVL LVKKDVESSLSSCNSSSTVSYLKKEDSVSQS SSDAGLGSDHESDTLTIDVSAISNLIRKHVSEA RLVEDIGHELTYVLPYEAAKEGAFVELFHEID DRLSDLGISSYGISETTLEEFLKVAEESGVDA ETSDGTLPARRNRRAFGDKQSCLRPFTEDDA ADPNDSDIDPESRETDLLSGMDGKGSYQVKG WKLTQQQFVALLWKRLLIARRSRKGFFAQIV

NO. of location control of the contr	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nuclecidid sequence where the contraspondid processor of the c			i .		beginning		D=Aspartic Acid, E=Glutamic Acid,
uence 9,496	nucl-	peptide			•	location	
uence 914 ng to first amino acid crisidue anino acid crisidue of peptide	cotide	seq-		_			I=Isoleucine, K=Lysine, L=Leucine,
minio acid residue of peptide residue of peptide residue of peptide sequence Particular and peptid	seq-	uence					
residue of peptide peptide peptide peptide peptide sequence Possible nucleotide deletion, "possible nucleotide deletion, "possible nucleotide deletion, "possible nucleotide deletion, "possible nucleotide deletion, "possible nucleotide deletion, "possible nucleotide insertion 1. PAYPYCLAIVESIL/PPFGKYPSLEI.QPWMY NEQYTEVSINDAPEDIGTIELI.NALTIC/PGFG TRCMEGON/PDITO/QAGEEWTTAPPYCDIT DLEYONONWTMQNPSPACQCSSINKIKMLEV CPPGAGGIPPPORKQNTADIQ, DUTCRNIKUM DLEYONONWTMQNPSPACQCSSINKIKMLEV CPPGAGGIPPORKQNTADIQ, DUTCRNIKUM POSSIBLE PRESIDENCE POSSIBLE PRESIDENCE POSSIBLE POSSIBLE PRESIDENCE POSSIBLE PRESIDENCE POSSIBLE PRESIDENCE PRESIDENCE POSSIBLE PRESIDENCE PRES	uence			914			
peptide sequence PayFyClaLyFsLivPpFGKYFSLELQPWMY		İ				,	
Sequence						sequence	
PAYFOCIAL/PSECNPERG YPSILELOPWMY NEQYTYS/NDAPEITOTH_LINALISDEGO TRCMEGNPIPDIPCOAGEEEWTTAPYPOTIM DLFONGOWTHONPSEA_OCCSSDKKKKMLPV CPPGAGGI_PPPORKONT ADIL ODLTORNISDY CPPGAGGI_PPORKONT ADIL ODLTORNISDY LYKTY_VOILAKSLKSKIW-VEPKYGGFSLEVS NTOALPPSOEVNDATY_COMKKKMLKLAKDSSA DPELNSLGRENTGLDTRNNVX_VEPHNING W HAISSELNVIDNALIKANI OKCENPSHYOTTAF NPENLIT KOQUSEV PAWTTSVDVLY_VEICVFA MSFVPASFYVHIORVYSKAKHLOFISGVXPYU VVLSSTNIP_VVLALLLLYQNSITTAMYPASFYK PSTAY_VVLSVLSTEDINGSVATTULLETION KLNNRDILKSVPLIPPHFCLGRCLIDMYRNQ AMADALERGERREVSPLSVDLYGRNLEAN AVEUVYPFLITVLQYPREIDEGVXPYU REKRARA VORLCVGIPPGECFGLLG VOKGACK SSTTKMLTGDITVTRODAFLNRYSILSNIHEV HQNMGVCPOFDATIELITGREFVEPFALROR VPEKEVGKVGEWAIRKLG_IXY_GEKYAGNY SGOKKRKLSTAMALIGOPTVVFLDEPTTOMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAMYNGFRCLOSVQHLKNRFOD GYTIVVRIAGSNPDLKVVQDFFGLAFFGSVYK SKHRNMM_VQLPSSLSARIFISLSQSKKRLH BEDYSVSQTTILDQYFVNFAKDQSDDDHLKDL SLHRQTYVVDVALVISTQDEKVKESYV TLLRGVSIIGTIGAGIPSKGAVQNTGSVGM SLTIMITVCQVLSLFGASYAQLOTTGSVGM SLTIMITVCQVLSLFGASYAQLOTTGSVGM VQLKGQTONKDAFSGRDSSTRLPLAFYYG MYAYAGWFYLNFYTEEVENPECTIPLACISM AVTIGTYVLTNAYSTITTMAEELLISNAVAVT PSERLLGRFSLAVIFVALSCFGSMGGYFAV UHLTMMLFSGOLDSLINFLSFARWLFGLA VAGLIVLRYKCPDMHRPFRVVLIPFLAFFTC LFMVALSLYSDPFTGIGFVTILTGYPAYYYLFU LHPLTMIMLFSGOLDSLINFLSFARWLFGLA VAGLIVLRYKCPGMHRPFRVVLIPFLAFFTC LFMVALSLYSDPFTGIGFVTILTGYPAYYYLFU LHPLTMIMLFSGOLDSLINFLSFARWLFGLA VAGLIVLRYKCPDMHRPFRVVLIPFLAFFTC LFMVALSLYSDPFTGIGFVTILTGYPAYYYLFU LHPLTMIMLFSGOLDSLINFLSFARWLFGLA VAGLIVLRYKCPDMHRPFRVVLIPFLAFFTC LFMVALSLYSDPFTGIGFVTILTGYPAYYYLFU LHPLTMIMLFSGOLDSLINFLSFARWLFGLA VAGLIVLRYKCPDMHRPFRVVLIPFLAFFTC LFMVALSLYSDPFTGIGFVTILTGYPAYYLFU LHPLTMIMLFSGOLDSLINFLSFARWLFGLA VAGLIVLRYKCPOMHRPFRVVLIPFAGSSGSE EAAGGEAGGDAGGWPGAGSSADQDDEGVVK PGNSLWPWDSVNRNILSRALTEMCLVTDVILSL VRKKFMTLDPVSQDALPFKQNPGTLQLISK KKSLAGAQULLKGAGRITKSVTENGCHENGL RDASSSGVRAVKISIGKAAPDIGLGTYN LFKRPLPFSKYGSFHWOTKLLAAQNFOKSGSFE EAAGTEGDAGWPGAGSSADQDDEGOVYR POSLUPMDSVNRNILSRALTBMCVLTQVDILLGE FAQLSSEAVQIKSQVPHIVVKNOISGPFSLO LSISLCKSSNDKKSGVFATIKGCCEBPILGYLET AGLISH LFKRPTATATATATATATATATATATATATATA							
NEQYTFVSNDAPEDITGITLELINALTXDPGFG TRCMEGRIPPIDITGOGREEWTTAPYOTIM DLFONGNWTHOMPSPACQCSSDKIKKMLPV CPROAGGLPPORKOMADILQULTGRISDY LVKTYVOJLAKSLKNKIWVNEPRYGGFSLOVS NTQALPPSGEVNDATADILQULTGRISDY LVKTYVOJLAKSLKNKIWVNEPRYGGFSLOVS NTQALPPSGEVNDATADILQULTGRISDY LVKTYVOJLAKSLKNKIWVNEPRYGGFSLOVS NTQALPPSGEVNDATADILQULTGRISDY LVKTYVOJLAKSLKNKIWVNEPRYGGFSLOVS NTQALPPSGEVNDATAGNIVAVWPNNKOW HAISSENIVINALIALNIQKGENPSHYGTTA MEPPLNLTKQQLSEVAPMTTSVVJVSCVIPA MEPPASTVYLIQRRYSKAKLQRISGVKPV YWLSNFYWDMCNYVVPATALVIHIPLOQKSY VSSTNLPVIALILLLLYGWSITIPLMYPASEVFK RETAYVOLTSVMLRIGINGSVATFVLJELFIDN KLINNINDILKSYFLIPPHELGRGLIDMYKNQ AMADALERGGENRYVSHSUDLVORNLFAM AVEGVVFFLITVLIQYRFFIRPRVNAKLSFLIN BEDEDVARERGRALLDGGONDILEIKELTIYY RRKKRAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGOTTYTRGOGONDILEIKELTIYY RRKKRAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGOTTYTRGOGONDILEIKELTIYY RRKKRAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGOTTYTRGOGONDILEIKELTIYY RRKKRAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGOTTYTRGOGONDILEIKELTIYY RRKKRAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGOTTVLTGOFALNINGED PKARRYLWCALSVYGGSVVLTSISMEC EALCTEMADMYNGRFRCLGSVQHLKNRFGD GYTTVLTGVNGASNPDLKKGGSVVLTSISMEC EALCTEMADMYNGRFRCLGSVQHLKNRFGD GYTTVLTGVNGASNPDLKKGGSVVLTSISMEC EALCTEMADMYNGRFRCLGSVQHLKNRFGD GYTTVLTGVNGASNPDLKKGGSVVLTSISMEC EALCTEMADMYNGRFRCLGSVQHLKNRFGD GYTTVLTGVNGASNPDLKKGGSVVLTSISMEC EALCTEMADMYNGRFRCLGSVQHLKNRFGD GYTTVLTGVNGASNPDLKKGGSVVLTSISMEC EALCTEMADMYNGRFRCLGSVQHLKNRFGD GYTTVLTGVNGASNPDLKKGGSVTANTSISMEC EALCTEMADMYNGRFRCLGSVQHLKNRFGD FKKRKRTADMIGFTGCTGTTTTKSMGGH FKKRKRTADMIGFTGCTGTTTTKSMGGH FKKRKRTADMIGFTGCTGTTTTKSMGGH FKKRTADMIGFTGCTGTTTTKSMGGH FKKRTADMIGFTGCTGTTTTKSMGGH TYTLEVFGQTATATANS LAFGRYLTSFTTGGGFTTTKSGGGH TYTLEVFGQTATATANS LAFGRYLTSFTTGGGFTTTTAGTATATANS LAFGRYLTSFTTGGGFTTTTAGTATATANS LAFGRYLTSFTTGGGFTTTAGTATATTTTAGTATATATATATATATATA		ļ			sequence		
TRÉMEGNPEDITECAGEEEUTTAPYPOTIM DLFONGWTMONPS ACQCSSDKKKKMLPV CPPGAGGLPPPQRKQNTADILQDLTGRNISDY LVXTYVOIJAKSLKNIKWVEPPXGPSLGVS NTQALPPSQEVNDATRQMKKHLKLAKDSSA DRFINSLGRFMTGLDTRNNVKVWFNNKGW HAISSFLWNDNALLAANLQKGENPSHYOTTAF NHPLNLTKQQLSEVAPMTTSVDVLVSICVIFA MSFVPASFVVRLIGERVSKAKHLGISGVKPVI YWLSNFVWDMCNTVVPATLVIIIFICRQGKSY VSSTNLPYLALLLLLYGWSITPLMYPASFVFK IPSTAYVVLTSVNLPIGINGSVATFVLELFTDN KLNNINDILSSVLLPIPHCLGRGLLDMYKNQ AMADALERGENRFVSPLSWDLVGRNLFAM AVEGVYFFLITVLIQYPFICRPRIVNAKLSLN DEDEDVRRERGRILDGGQNDILEIKELTKIY REKRKPAVDRICVGIPPGEGFGLLGVKGAKK SSTFKMLTGDITVTRGDAFLNRNSILSNIHEV HONMGVCPQFDATIELITGRHVPFFFALIKG VPEKEVCKVGEWARKLGLVKYGEKYAGNY SGGKKRLSTAMALIGGPTVVFLDEFTGMD PKARRFLWNCALSVVKEGRSVVLTSISMEGC GATTAWAMTVGRFCLGSVQHLKVKSFOR SGHKRKLSTAMALIGGPTVVFLDEFTGMD PKARRFLWNCALSVVKEGRSVVLTSISMEGC GATTAWAMTVGRFCLGSVQHLKVKSFOR SCHRMAMTVGRFCLGSVQHLKVKSFOR SCHRMAMTVGRFCLGSVQHLKVKSFOR SCHRMAMTVGRFCLGSVQHLKVKSFOR SCHRMAMTVGRFCLGSVQHLKVKSFOR SCHRMAMTVGRFCLGSVQHLKVKSFOR SCHRMAMTVGRFCLGSVQHLKVKSFOR SCHRMAMTVGRFCLGSVQHLKVKSFOR SCHRMAMTVGRFCLGSVQHLKVKSFOR SCHRMAMTVGRFCLGSVQHLKVKSFOR SCHRMAMTVGRFCLGSVQHLKVKSFOR SCHRMAMTVGRFCLGSVGHLKVKSFOR SCHRMAMTVGRFCLGSVGHLKVKSFOR SCHRMAMTVGRFCLGSVGHLKVKSFOR SCHRMAMTVGRFCLGSVGHLKVKSFOR SCHRMAMTVGRFCLGSVGHLKVKSFOR SCHRMAMTVGRFCLGSVGHLKVKSFOR SCHRMAMTVGRFCLGSVGHLKVKSFOR SCHRMAMTVGRFCLGSVGHLKVKSFOR SCHRMAMTVGRFCLGSVGHLKVKSFOR SCHRMAMTVGRFCLGSVGHLKVKTSTGSVGN SCHRMAMTVGRFCLGSVGHLKVKTSTGSVGN SCHRMAMTVGRFCLGSVGHLKVKTSTGSVGN SCHRMAMTVGRFCLGSVGHLKVTSTGSVGN SCHRMAMTVGRFCLGSVGHLKVTSTGSVGN SCHRMAMTVGRFCLGSVGHLKGCGFGETUL SCHRMAMTVGRFCLGSVGHLKGCGFGETUL SCHRMAMTVGRFCLGGGGGGE SCHRMAMTVGRFCLGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	i	ŀ	ľ		! 		
DIFONGNITMONPSPACQCSSDKIKKMUS POPPAGGGI-PPORKONTADLOQLIGRNISDY LVKTYVQIJAKSLKNKIWVNEPRYGGPS12/SV NTQALPPSQEVIDATKQMKKHKLAKDSSA DREINSLGREMTGILDTRNIVKVWENNKGE HAISSENVIDATIKAMI, QKESPSHYGITAF NIPILNITKQQLSEVAPMTISVDVI, VSICVIFA MSFVPASFVVFI. IGBERVSKAKILGFGKVPVI YWLSNFVWDMCNYVVPATLVIIIPICPOQKSY VSSTINPVI.LILLILLYGWISTRLMYPASFVFK PSTAYVVLTSVNLFIGINGSVATFVLLIFTIDN KLANINDILKSVFLIFPIFFCLGRGLIDMYKNQ AMADALERGGENFYSPLSWIDVGRNLFAM AVEGVVFFITVLIQYRFFIRRPFVNAKLSPLI DEDEDVRERQRILLYGGGQNDILEKEITKIY RRKRFPA VDRICVGIPPGECFGLLGVVGAGK STIFKMI.TGDITVTRGOFAFINRSILSNIHEV HQNMGYCPGPAITELLTGREHVEFFALLEN PSTAYVVLTSVNLFIGINGSVATFVGEKYAGNY SGGNKRILSTAMALIGGPFVVFLDEFTTGMW FKARRFLWCALSVKEGRSVVYGEKYAGNY SGGNKRILSTAMALIGGPFVVFLDEFTTGMW SGGNKRILSTAMALIGGPVVFLDEFTTGMW FKARRFLWCALSVKEGRSVVYKGERSVVTSHSMEEC EALCTRNAMINVNGFRCLGSVQHLKNRFGO GYTIVVRIAGSNPDLKPVQDFFGLAFGSVFK EKHRMLQVQU.PSSLSSLAAIFSILSQSKKRIL EDYSYSYQTTLDQVFVFNFAKQSDDDHLKDL BEYSYSYQTTLDQVFVFNFAKQSDDDHLKDL SHNNQTVVDVAVLTSFLOBEKVESSYV TLKGSGGGIFSKGVGLYGFSVGW SLHWINGVGLSGLGSSKGRGKGRFQGEKYQLKKKL VLGSWIGNTIGAGIFSKGCGLYGTGSVGW SLTIWTVCGVLSJEGALSYAELGTTIKKSGGH YQGWNGRLPSLGRHERPCVFFVSTISKGG YLQGNVNGRLPSLGRHERPCVFFVSTISKGG YLQGNVNGRLPSLGRHERPCVFFVSTISKGG YLQGNVNGRLPSLGRHERPCVFFVSTISKGG YLQGNVNGRLPSLGRHERPCVFFTALILIVGV WQLIKGGTQNFKDAFSGRDSSITRLPLAFYY MQLIKGGTQNFKDAFSGRDSSITRLPLAFYY MQLIKGGTGNFKDAFSGRDSSITRLPLAFYY MQLIKGGTGNFKDAFSGRDSSITRLPLAFYY SERLILGNSLAVPTEEVEPPEKALINGN AVTIGYYLTNAVFTTINAEELLISNAVAY SRLFYVASREGHILPELSMITTAVGTYVM UNSMSVSWAGNIGHTJCKLITAPLGISM AVTIGYYLTNAVFTTINAEELLISNAVAY SRLFYVASREGHILPELSMITTAVGTYVH UNSKYRWRIMSSKTRNTQIILEVYPEDEXL RGGGLLSSRLSAKGPURTSFTGGWGVLP LADAASMSGVRAVRISIESACKQMFEVILIFEGWGLY LADAASMSGVRAVRISIESACKQMFEVILIFEGWGLY POPSTUREAGRAFIKTSVTRQENKLQ RDFNSELIRLRQHWKLRKVGDKILGDLSYRS AGSLFFHGTFEVKNTDLDLDKAPLCKEI FAQLSREAVQIKSQVPHIVKNOJLGCEI FAQLSREAVQIKSQVPHIVKNOJLGCEI FAQLSREAVQIKSQVPHIVKNOJLGCEI FAQLSREAVQIKSQVPHIVKNOJLGCEI FAQLSREAVQIKSQVPHIVKNOJLGCEI FAQLSREAVGIKSGVPHIVKNOJLGCEI FAQLSREAVGIKSGVFHITVKNOJLGCEI FAQLSREAVGIKSGVFHITVKNOJLGCEI FAGLSREAVGIKSGVFHITVKNOJ		1					
CPPGAGGIPPORKONTADILODLIGNINSON	ļ						
IVKTTVQILAKSLKNKIWVERTYGGESLOV NTQALPPSQEVIDATKQMKKHKLAKDSSA DRFLNSLGRPMTGLDTRNVKVWPINNKQN HAISSI-NVINNALIRANL QKEBPSHYGITAF MFPLNLTKQQLSEVAPMTTSVDVLVSICVIFA MSFVPASFVVPLIQERVSKAKHLQFISGVKPVI VINLSNFVWDMCNYVVPATLVIIIFLCPQKSY VYSSTNLPVLALLLLLYGWSITFLMYPASFVK RPSTATVVLTSVNLFIGINGSVATFVLELFTDN KLNNIDDILKSVFLIPFHCLGRGLIDMYKNQ AMADALERFGENFYSPLSWDLVGRNLFAM AVEGVYFFLITVLQVRFFIRPRPVNAKLSPLN DEDEDVRERQRIDLGGGQNDILEKLTIY RRKRFPAVDRICVGIPFGECFGLLGVNGAGK STEKMLTGDITYTRGDAFLNRSILSNIHEV HQNMGYCPFDATTELLTGREHVEFPALLEG HQNMGYCPFDATTELLTGREHVEFPALLEG PVEKEVGKVGEWARKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPVVFLDEFTTGM PKARRFLWNCALSVYKEGRSVYLTSISMEC EALCTRMAIMVNGRFRCLGSVQHLKNRFDD GYTTVYRLGSNPDLGFVCYSTISSMEC EALCTRMAIMVNGRFRCLGSVQHLKNRFDD GYTTVYRLGSNPDLGFKVGEKYAGNY SGGNKRKLSTAMALIGGPVVFLDEFTTGM PKARRFLWNCALSVYKEGRSVYLTSISMEC EALCTRMAIMVNGRFRCLGSVQHLKNRFDD GYTTVYRLGSNPDLGFKVGEKYAGNY SGGNKRKLSTAMALIGGPVVFLDEFTTGM FRANKEN			İ		}		
NTQALPPŠQEVNDATKQMKKHLKLAKDSSA BEFINSLGRPMGLDTRNNVKVWPNNGGW HAISSFLANDNALIRAAILQKGENPSHYGITAF NHPLNLTKQQLSEVAPMITSVDVLVSICVITAF NHPLNLTKQQLSEVAPMITSVDVLVSICVITAF NHPLNLTKQQLSEVAPMITSVDVLVSICVITAF MSFVPASFVVHLQERVSKAKILQFISGVKPVI YWLSNFVWDMCNYVVPATLVIIIFICPQKSKY IPSTAYVVLTSVNLFIGINGSVATFVLELTION KLNNIDDILKSVLIPPFICLGRGUDMVKNQ AMADALERFGENFVSPLSWDLVGRNLFAM AVEGVVFLITVLIQFSPTGVAGK STFKMLTGDTTVTRGDAFLNRNSILSNIHEV HQNMGVCPQFDAITELLTGREHVEFFALIRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGFPVVFLDEPTTGMS PKARFFLWNCALSVYKGEGSVVTYSTISMEEC EALCTRNAMINVNGRFRCLGSVQHLKNRFGD GYTTVYNIAGSNPDLKPVQDFFGLAFGSVPK EKHRMLQVQLPSSLSSLARIFSILSQSKKRH EEPYSVSQTTLDQVYVFAKADGSDDHLKDL BEFYSVGTTLTQVYVFAKADGSDDHLKDL SLHKNQTVVDVAVLTSFLODEKVKESYV TLLBGVSUGTTIGVVVFNFAKOSDDHLKDL SLHKNQTVVDVAVLTSFLODEKVKESYV TLLBGVSUGTTIGVTVFNFAKOSTISKGG YLQGNVNGLPSLGRKEPPGQEKYQLKREV TLLBGVSUGTTIGVTVFNVAVLSFRALIFIRATAVIT VANASVSWSARQIFISFKGGVLQNTGSVGM SLTIWTVCOVLSLFGALSVALGGTTIKKSGGH YTYTLEVFQH_PAFVRVWVELLIRRAATAVIT VM VLNSMSVSWSARQIFISTFCKLTALIIIIVPGW MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTENAFELLSNAVAVIT FSEELLGNPSLAVPIFVALSCFGSNMGGVFAT SEELLGNPSLAVPIFVALSCFGSNMGGVFATE UFFTMMLSSQLAVTTINAESCLISNAVAVIT FSEELLGNPSLAVPIFVALSCFGSNMGGVFATE UFFTMMLSSGLAVTTINAESCRUPFFGGLA VAGLIYVLTNVAYDISHALSNAVAVIT FSEELLGNPSLAVPIFVALSCFGSNMGGVFATE UFFTMMLSSGLAVFTINAESCRUPFFGGLA VAGLIYVLTNVAYDISHALSNAVAVIT FSEELLGNPSLAVPIFVALSCFGSNMGGVFATE UFFTMMLSGOLDSLLAFT FSEELLGNPSLAVPIFVALSCFGSNMGGVFATE TRANSLAVRISHSACREGVHEVGLU RDFNSELLARQHMFCLRKVGDKLIGDLSYS SRLFYVASREGHLEGLISMHIVAKTRYDPAVYU LIPITMMLSGOLDSLLAFT FSEELLGNPSLAVPIFVALSCFGSNMGGVFATE TRANSLAVRISHSACREGVQHEVGLU RDFNSELLARQHMFCLRKVGDKLIGDLSYS SRLFYVASREGHLEGNETHOPAVITOLLSIK KKSLAGAQULLKGAERLTKSVTENQENKLQ RDFNSELLARQHMFCLRKVGDKLIGDLSYS AGSLFFHGTFEVKNTOLDLDKKREPDYCPL DVQPISDLEGSAYHWSLORQAPDIGDLGTVU LFKPPLPKSKFGSPHWOTTLCAAGNVALLCKEI FAQLSREAVQIKSQOYPHIVVKNONISGPFFSUC LFKPLPKSKFGSFFWOTTLCAAGNVALLCKEI FAQLSREAVQIKSQOYPHIVVKNONISGPFFSUC LSISLCHSSNDKKSOKGRATEKQCFGPHLTYFLQ							
DREINSLGREMTGLDTRNNVKVWENNKGW HAISSELNVINALIRANIQKESPHYOTITAF NHPLNITKQQLSEVAPMTTSVDVLVSICVFA MSEVPASSVVPLIQERVSKAKHLQFISGVKPVI YWLSNPVWDMCNYVVPATLVIIIE(CQKSY YSSTNLPVLALILLILYGWSIPPLMYPASFVFK PSTAYVVLTSVNLFIGINGSVAFVILLELFTDN KLNNINDLKSVFLIFPHPCLGRGLIDMYKNQ AMADALERRGENFVSSLEWUGRNEAM AVEGVVFFLITVLIQYRFIRRPRVNAKLSPIN DEDEDVRERGRILDGGGGNDLIEKELTKIY RRKRKPAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGDITVTRODAFLNRNSILSNIHEV HQMMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKGEKYAGNV SGGNKRKLSTAMALIGGPPVVFLDEPTITGMD PRARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSPDLKPVQDFGLAFFGDVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH IEDVSSQTTLDQVFVNFAKDQSDDDHLKDL SLIKNRQTVVDVAVLTSRLQDEKVKESYV TLRGVSIIGTIIGAGFISPKGVLQNTGSVGM SLITIWTVCGVLSFGALSVAGKENTAVIS LAFGRYLLEPFFCVFTMYKRPVVSTISKGG YQGNNGKLSTAMALIGFPFCVCKGKYGNV VINSMSVSNARIOFILTFCKLTAILIIVPGV MQLISGTQNFKAFSGROSDCKVQLKRKV TLLRGVSIIGTIIGAGFISPKGVLQNTGSVGM SLITIWTVCGVLSFGALSVAGKTAVISTEAGA VVTIJELFYGPLFAFVRVWVELLIBPAATAVIS LAFGRYLLEPFFCCEPELAKITAVGTTVM VLNSMSVSWSARIOFILTFCKLTAILIIVPGV MQLISGTQNFKAFSGROSDCSTETKKSGGH YYTILEVFGPLFAFVRVWVELLIBPAATAVIS LAFGRYLLEPFFCCEPELAKITAVGTTVM VLNSMSVSWSARIOFILTFCKLTAILIIVPGV MQLISGTQNFKAFSGROSDLSLLNSAAVFFLACISM AVTIGVYLTNVAYFTTINAELLLSNAAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREOHLPELLSMHVRKHTPLAVTV LIPPLTMIMLFSGDLDSSLLNFLTVAAYVTFITEVENPEKTTELACISM AVTIGVYLTNVAYFTTINAECLLSNAAVFF FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREOHLPELLSMHVRKKTPLAVTV GMYAAGGVFLOPPLTVILGTURATAVFT WDKKPRWFRIMSEKITRTQIILEVVPEEDKL RGGGSLRCVLGKLIGGLGCGSRCVGPPEGLL RFRGCGLLSSRLSAGRPLRTSFFGSWQLPP LADAASMSGVAVVRIISSAGCRGQVFPVLSID GTETYLPPLSMSGNLARLAGRIDFSQGSGSEE EBAAGTEGDAGOBPCAGSSADQDEGGVVX FQPSLWPNDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPFKQNPGTILGJISK KSKSLAGAAQILLKGABLITKSVTENQENKLQ RSFFNSELLRLRQWKKLRKVGDKLGDGTVN LFKRPLPKSKFPSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHTVVKNQISLGGTVN LFKRPLPKSKFPSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHTVVKNQISLGGTVN LFKRPLPKSKFPSPHWQTKLEAAQNVLLCKEI FAQLSCREAVQIKSQVPHTVVKNQISLGGTPHLYVLE FAQLSCREAVQIKSQVPHTVVKNQISLGGTPHLYVLE FAQLSCRE						:	
HAISSEINVRNALIRANI, OKGENPSHYGITAF NEPILAIT, TO, QISEVAPMITY DVISICUPE MSPVRASPVYELIQERVSKAKHLOPISGVKPVI YWLSNPVWDMCNYVVPATLVIIIFICFQQKSY VSSTNLPVLALLLILYGWSTPILMYPASSVKR IPSTAYVVLTSVNLPIGINGSVATFVLLEITIN KLNNNDILKSVFLIPHIPCLGRGILDMVKNQ AMADALERGENRIVSPLSWDLVGRNILAM AVEGVVFELITVLQYRFIRERPRVNAKLSPLN DEDEDVRERQRILDGGGQNDILEIKELIKIY RKRKPAVDRICVGIPPGECFGLIGVNGAGK SSTFKMLTGDITVTRGDAFLININSILSNIHEV HQNMGVCPQFDAITELLTGREUPFFALLRG VPEKEVGRVGEWARKLGLUVKYGEKVAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PRARRFUNVALSTVKEGSVQHLKNRFGD GYTIVVRIAGSPDLKPVQDFEVFFALLRG GYTIVVRIAGSPDLKPVQDFSDDDHLKDL EDVSSQTILDQVFVNFAKDDDDHLKDL SEHRINMLQYQLPSSLSSLARIFSILSQSKKRLH EDVSSQTILDQVFVNFAKDDDDHLKDL SHKNGTVVDVAVLTSFLQDEKVKESYV 11LRGVSIIGTILGAGFISPKGVQLDATSVCM SLIHKNGTVVDVAVLTSFLQDEKVKESYV 11LRGVSIIGTILGAGFISPKGVQLDATSVCM SLIHKNGTVVDVASPLKPVSTISKGG YTQGNNGRLPSLGNKEPFQQEKVQLKRKV TILLRGVSIIGTILGAGFISPKGVQLDATSSVCM SLIHKVCOVLSLFGALSYAELGTTIKKSGGH YYTHEFFFQCEBELAKHTAVGTVVM VLNSMSVSWSARQIFLTFCKLTALLITYPGV MQLIKGQTQNFAJASCGSSNGGYFAV SKLTYVASRGGHLPELISMINATAVGTTVM VLNSMSVSWSARQIFLTFCKLTALLITYPGV MQLIKGQTQNFAJASCGSSNGGGFAV SKLTYVASRGGHLPELISMINATT FSERLLGNFSLAVPIFVALSCGSSNGGYFAV SKLTYVASRGGHLPELISMINATT FSERLLGNFSLAVPIFVALSCGSSNGGYFAV SKLTYVASRGGHLPELISMINATT FSERLLGNFSLAVPIFVALSCGSSNGGYFAV SKLTYVASRGGHLPELISMINATT FSERLLGNFSLAVPIFVALSCGSNNGGYFAV SKLTYVASRGGHLPELISMINATT FSERLLGNFSLAVPIFVALSCGSNNGGYFAV SKLTYVASRGGHLPELISMINATT FSERLGNFSLAVPIFVALSCGSNNGGYFAV SKLTYVASRGGHLPELISMINTOLLOKFTPEDKOLL RRGCGLLSSRLSAGKPPLRTSFFGSWQLPP LADAASMSGVAVRIISBAGERGVYHEVGLD GTETYLPPLSMSQNLARLAQRIDPSQGSGSEE EBAAGTEGDAGDWPGAGSSADDDEGGVVX FOPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPSQDALPFROPTQLLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFPSSLERLRGWKKRKVGDKLIGDTVN LFKRPLKSKFSPHWOYKLEAAQNVILLCKEI FAQLSERAVQIKSQVPHTVVKNQUISCPFFSQL LISKCHSLDKKSQKAPHTVKNQUISCPFFSQL LISKCHSSINKKSVGKAAPDIGDLISKYR AGSIFPHHGTTERHVKNKONGUISCPFFSQL LISKCHSSINKKSVGKAAPDIGDLISKYR FAGLSERAVQIKSQVPHTVVKNQUISCPFFSQL LISKCHSSINKKSVGKAARDVILLCKEI FAQLSERAVQIKSQVPHTVVKNQUISCPFFSLQ LISKCHSSINKSGKAPHOVILLCKEI FAQLSERAVQIK						:	
MIFLINITRQQLSEVAPMTTSVDVLVSICVIPA MSFVPASFVVPLIQERVSKAKHLQFISGVKPVI YWLSNPVMOMCNYVVPATLVIIIFICQKSY VSSTNLPVLALLLLI YGWSITPLMYPASFVFK IPSTAYVLTSVNLPIGINGSVATFVLELFTIDN KLNNINDILKSVFLIFPHPCLGRGLIDMYKNQ AMADALERGERFRVSPLSUDVGRNLFAM AVEGVVFFLITVLIQYRFIRFRPVNAKLSPIN DEDEDVRERGRILDGGGGNDLIEIKELTKIY RKRKPAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGDTTVTRGDAFLNSIS.SINHEV HQMMGYCPQFDAITELLTGREHVEFFALLEG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVYPLDEPTTOMD PKARRFLWNCALSVVKEGRSVVLTSHSMEC EALCTRMAMNYGRFRCLGSVCHLKNRFGI GYTTVVRIAGSNPDLKPVQDFFGLAFPGSVFK EKHRNMLQVQLPSSLSSLARFIS,GSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDPHLKDL SLHKNQTVVDVAVLTSFLQDFEVALSKGSV YLQGNVNGRLFSLGNKEPFQCEKVQLKRKV TILLRGVSINGTINGAGFISPKSVLQNTGSVGKR YLQGNVNGRLFSLGNKEPFQCEKVQLKRKV TILLRGVSINGTINGAGFISPKOLVQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTKKSGGH YTYTILEVFGPLFAFFVXVWELLIIPPAATAVIS LAFGRYILEFFFIQCEFPLAIKLITAVGITVVM VINSMSVSWARNIOFLTFECLTAILLIIVPGV MQLIKGGTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCGSNNIGGFAV SKLFVVASRGEHLPELSMHVRKHTPLPAVIV LIPHTMMLFSGDLDSLLNFLSFARWLFIGLA VAGGIYLRYKCPDMHFPFRVLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYJFII WKKPRWFNIMBEKTITRLQILEVYPEDWL VAGGIYLRYKCPDMHFPRVLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYJFII WKKPRWFNIMBEKTITRLQILEVYPEDWL RFRGCOLLSSRLSAGKPPLRTSFFGSWQLIP LADAASMSGVRAVRSIESBAGNQLDEEGVVK FOPSLWPUDSVRNNLRSALTEMCVLYDVLSI VRDKKFPMLTPPSSQDALPFRVGPTULJSK VRDKKFPMLTPPSSQDALPFRSQCSEE EEAAGTEGDAQEWGAGSSAGEEQVHEVGLD GTETYLPPLSSMONLARLAQRIDFSQGSSEE EEAAGTEGDAQEWGAGSSADAQDDEEGVVK FOPSLWPUDSVRNNLRSALTEMCVLYDVLSI VRDKKFPMLTPDYSQDALPFRYDOTIOLISK KKSLAGAAQILLKGARRLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKLODLSYRS AGSIFPHGTFEVIKNTDLDLDKKPEDYCCL DVQIPSDLEGSAYIKVSIQKQAPDIGDLJCKTPETYCL LFKRPLKSKFPGSPHWVYLKAEAAQNVLLCKEI FAQLSERAVQUKSQVPHTVVKNQUISCPFFSLQ LISICHSSDNKKSQKAATRICGDLGTVN LFKRPLKKSKPGEPHHQYVLCEPHLTYVLE LEGERAATGEDAGEPHLTYCHE FAQLSCREPHTYCHELFTYCHELTYCHELFTYCH LFKRPLKSKFPGEPHHQVLLCEPHLTYLLE FAQLSERAVQUKSQVPHTVVKNQUISCPFFSLQ LISICHSSDNKKKSQKAATRICGDLGTVN LFKRPLKSKFPCEPHLTYVLLE							
MSFYPASFVÝRLIQREVŠKAHLQFISGVKPNI YWLSNPVUMOKONYVPATLVIIIPICFQQKSY YSSTNLPVLALLLLI YGWSITPLMYPASFVFK IPSTAYVVLTSVALPIGINGSVATFVLELFTDN KLNNINDILKSVFLIPPHCLGGGLIDMYKNQ AMADALERGENRFVSRLSWDLVGRNILAM AVEGVVFRLIVTU,QYRFIPRPVNAKLSPLN DEDEDVRERQRILDGGGQNDILEIKEITKIY RRKRFA VDRICVGIPPGECFGLIGVNGAGK SSTFKMLTGDITVTRGDAFLNRNSISNIHEV HQMMGYCPQFDAITELLTGREFVEFFALLRG VPEKEVGKVGEWARKI,GLVKYGGKYAGNY SGGNKRKLSTAMALIGGPPVLDEPTTCMD PKARRT WNCALSVYKEGRSVVLTSHSMEEC EALCTRMAIMVNGFRCLGSVOHLNRFGD GYTVVRIAGSPPLLKPVQDFFDAFPGSVFK EHRINNLQVQLFSSLSSLARIFSILSQSKKRLH EDVSVSQTILDQVFVNFAKDDDFHLKDI. SLHKNGTVDVAVLTSFLQDEKVKESYV 11LGVSIIGGTILGAGFISFRGVLQNTGSVGM SLTIWTVCQVLSLFGALSYAELGTTIKKSGGH YTYLEVFGPLAFVRVWVLLIRPAATAVIS LAFGRYLLEPFFIQCEIPELAKLITAVGITVVM VLNSMSVSWSARQIFILTFCKLTALLITVPGV MQLIKGQTQNFAAFSGRDSSTIRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIFLALGISM AVVIGYLTNWYFTTINAEGLLISNAVAVT FSERLLGNISLAVPIFVALSCGSNNIGGFAV SKLFYVASREGHLPELISMINVKHTPLPAVIV LIPHTMIMIFSGDLDSLLNFLSFRAWLFGLA VAGIJYLRYKCPDMHFPETVPLFPPLEDSKLI RRGCGLLSSRLSAGKPPLRTSFFGSWGVPP LADAASMSCVRAVRISESACRGVYHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EAAGTEGDAGEWPGAGSSAGCHQTHESGNLQ ROFPSELLRAGHWKRKYCDGRIGTVN LFKRRJPKSKFPGDEWPGNLQ ROFPSELLRAGHWKRKYCDGRIGTVN LFKRRJPKSKFPGSPHWOYLLGKAQAPDIGDLLKYRS VERKKFMTLDPSQDALPFROQLEGIVN LFKRRJPKSKFPGSPHWOYLLGKAQAPDIGDLISYRS AGSIPPHGTTEVIKNTDLDLDKKIPEDYCL DVQIPSDLEGSAYIKVSIQKQAPDIGDLISYRS AGSIPPHGTTEVIKNTDLDLDKKIPEDYCL DVQIPSDLEGSAYIKVSIQKQAPDIGDLISYRS AGSIPPHGTTEVIKNTDLDLDKKIPEDYCL DVQIPSDLEGSAYIKVSIQKQAPDIGDLISYRS AGSIPPHGTTEVIKNTDLDLDKKIPEDYCL DVQIPSDLEGSAYIKVSIQKQAPDIGDLISYRS AGSIPPHGTTEVIKNTDLDLDKKIPEDYCL DVQIPSDLEGSAYIKVSIQKQAPDIGDLISYRS AGSIPPHGTTEVIKNTDLDLDKKIPEDYCL DVQIPSDLEGSAYIKVSIQKQAPDIGDLISYRS AGSIPPHGTTEVIKNTDLDLDKKIPEDYCL DVQIPSDLEGSAYIKVSIQKQAPDIGDLISYRS AGSIPPHGTTEVIKNTDLDLDKKIPEDYCL DVQIPSDLEGSAYIKVSIQKQAPDIGDLISYRS AGSIPPHGTTEVIKNTDLDLDKKIPEDYCCL DVQIPSDLEGSAYIKVSIQKQAPDIGDLISYRS AGSIPPHGTTEVIKNTDLDLDKKIPEDYCCL DVQIPSDLEGSAYIKVSIQKQAPDIGDLISTYRS AGSIPPHGTTEVIKNTDLDLDKKIPEDYCCL DVQIPSDLEGSAYIKVSIQKQAPDIGDLISTYRS AGSIPPHGTTEVIKNTDLDLDKKIPEDYCC]	ļ			
YWLSNFWDMCNYVPATLVIIIFICPQKSSY VSSTNLPVLALILLLYWSSTIPIMYPASFVFK PSTAYVVLTSVALFIGINGSVATFVLELFTDN KLANINDIL SKYPLIFPIBLGRGLIDMYKNQ AMADALERFGENRFVSPLSWDLVGRNLFAM AVBGVVFFLITVLIQVRFFIRPRYVNAKLSPLN DEDEDVRRERQRILDGGQONDILEIKELTKIY RREKRFAVDRICVGIPPGECFGLLGVNGAGK SSTFMLTGDTTVIRGDALNRNSILSNIHEV HQNMGYCPGPAITBLLTGREHVEFFALLRG VPEKEVGKVGEWARRALGVYGEKVAGHY SGONKRKLSTAMALIGGPVVVFLDEPTTOMD PKARFLWNCALSVVKGERVVLITSHSMEEC EALCTEMAMVNGFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQVQLPSSLSSLARFISILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDHLKDL PSTATLDQLFVVFNFAKDQSDDHLKDL SHKNQTVDVAVLTSHQDEKVKESYV TLRGVSIIIGTIIGAGFISFRGVLQNTGSVGM SLTIWTVCQVLSLFGALSYAELGTTIKKSGGH YTYLLEVFGELPELAIKLITAVGTVVM VLNSMSVSWARQIGHLFGLSTAVAIS LAFGRYLLEPFFIQEEPELAIKLITAVGTVVM VLNSMSVSWARQIGHLFDPSTYPLACISM AIVTIGYVLINVAYFTTINAEELLISNAVAVT PSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLEELISNHTVRKHTPLPAVIV LHELTMIMLFSGDLDSLLNRLSFSTC LFMVALSLVSPFFSTGGFVITLTGUFAYYLFII WKKPRWFRMSKENTRTLQIILEVVPEDKIL WKKPRWFRMSKENTRTLQIILEVVPEDKIL WKKPRWFRMSKENTRTLQIILEVVPEDKIL WKKPRWFRMSKENTRTLQIILEVVPEDKIL WKKPRWFRMSKENTRTLQIILEVVPEDKIL WKKPRWFRMSKENTRTLQIILEVVPEDKIL WKKPRWFRMSKENTRTLQIILEVVPEDKIL RHRGCGLLSRLSAGKPPLRTSFTGSWGVLPP LADAASMSGVRAVRISISSACEKQVHEVGLD GTETTLPFLSMSQNLARLAGRIDFSGGSGEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VROKKFMTLDPVSQDALDLSYRS AGSLFPHIGTFEVINNTDLDLDKKPPEDYCPL DVQIPSDLEGSAYIKVSIQKQADDIGLGTVNL RRFRCGLLSRLSAGKVPHIVKLKVGLDLSYRS AGSLFPHIGTFEVINNTDLDLDKKPPEDYCPL DVQIPSDLEGSAYIKVSIQKQADDIGLGTVNL KKRELLRCHWKLRKVGOTEDHLYVLE LISCHESSBNKKSQKFAFREGGFHOUTLCKEI FAQLSREAVQUIKSQVPHIVVKNQIISSPFSLQ LISISCHESSBNKKSQKFAFREGGFHOUTLCKEI FAQLSREAVQUIKSQVPHIVVKNQIISSPFSLQ LISISCHESSBNKKSQKFAFREGGFHOUTLCKEI FAQLSREAVQUIKSQVPHIVVKNQIISSPFSLQ LISISCHESSBNKKSSQKFAFREGGFHOUTLCKEI FAQLSREAVQUIKSQVPHIVVKNQIISSPFSLQ LISISCHESSBNKKSSQKFAFREGGFHOUTLCKEI FAQLSREAVQUIKSQVPHIVVKNQIISSPFSLQ LISISCHESSBNKKSSQKFAFREGGFHOUTLCKEI FAQLSREAVQUIKSQVPHIVVKNQIISSPFSLQ LISISCHESSBNKKSSQKFAFREGAAQNVLLCKEI FAQLSREAVQUIKSQVPHIVVKNQIISSPFSLQ LISISCHESSBNKKSSQKFAFREGAAQNV							
VSSTNLPVLALLILLLYGWSITPLMYPASFVPK PPSTAYVVLTSVNLFIGINGSVATFVLEIFTDN KLNNINDILKSVFLIFPHFCLGRGLIDMVKNQ AMADALERGOENFYSPLSWDLVGGNLFAM AVEGVVPFLITVLIQVRFIPRPRVNAKLSPLN DEDEDVRRERQRILDGGQONDILEIELTKIY RRRKFRAVDRICVGIPPGEGFGLLGVNGAGK SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV HQMMGYCPGPAJTELLTGREVTEFFALLEG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEFTTGMD PKARRLWNCALSVVKEGRSVVLTSHSMEEC EALCTTMAMINVNGRFRCLGSVQHLKNRFGD GYTTVVRILAGSNPDLKFVQFGLAPFGSSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH EDDYSVSQTTLDQVFVNFAQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPFCVPTMYRKFVSTISKGG YTLLEVGGNINGTPLGKVFTSTSKGGH YTYLLEVFGPLPAFVRWVELLIRPAATAVIS LAFGRYLLEPFFIQCEIPELAKLITAVGITVM VLNSMSVSWSARIQFILTFCKLTAILLIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAVFTTINAEELLLSNAVATI PSERILGNFSLAVPIFVALSGSMGFAV SRLFYVASREGHLPELSSKHWRKHFILA VAGLYYLRYKCPDMIRPFKVPLFFALFSFTC LFMVALSILVSDFFSTGIGFVITLTGVAYYLFIT 978 2328 A 8261 2 2165 RGGSIRCVLGKLIGQLLCFQSERCYPFEGLL LFMVALSILVSDFFSTGIGFVITLTGVAYYLFIT MUKKFRWFRIMSEKITTRLQIILEVVPEEDKL RGGGLISSRLSAGKPLTSFFGSWGVLP LADAASMSGVRAVRISIESACECQVHEVGLD GTETYLPPLSMSQGSSEE EBAAGTEGDAQEWPGAGSSADQDDEGGVVK FQPSLWPWSVRNNLRSALTEMCVLTDVLISL VRNKKFMTLDPVSODALPKONPOTICULISK KKSLAGAQUILKGAGRLTKSVTENGENICQ RPDFNSELLRIRLQHWKLIKAVGRLGDLSYRS AGSLPPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGGIGTVN LKRRIPLFSKRFGSFHWQTIKVORLIGDLSYRS AGSLPPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGGIGTVN LKRRIPLFSKRFGSFHWQTIKLAQNIVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISSPPSLQ LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE LISLGLESSNOKKSQKFATRISCOCPEDHLYVLE							
PSTAYVVLTSVNLIFIGINGSVATFYLELFTDN KLANNIDILKSVFLIFPHERGOLIDMYKNQ AMADALERFGENRFVSPLSWDLVGRNLFAM AVEGVVFFLITVLIQVRFFIRPRYVNAKLSPLN DEDEDVRRERQRILDGGGONDILEIKELTKIY RREKRFAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGIDTVIRGDALNRNSILSNIHEV HQMMGYCPGPAITBLLTGRBHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTISMMEC EALCTEMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKEGSVQHLKNRFGD GYTIVVRIAGSNPDLKESTV EEKHRNMLQYQLPSSLSLAFISLSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV PGSTISPSLCFIPPCVPTHAVFRVVSTISKGG YLQGNVNGRLPSLGNEEPPGGEKVQLKRKV TLLRGSSIIGTIGAGGIFSKGVLQNTGSVGM SLTHVTVCQVLSLFGALSYAELGTTIKKSGGH YTYLLEVFGPLRAFVRVWVELLIRFAATAVIS LAFGRYLLEPFFQCEPPELAIKLITAVGITVVM VLNSMSVSWSARIQFLTFCKLTALLIIVPGV MQLKGGTGONFKDAFSGRSTIRLPLAFYYG MYAYAGWPYLNFVTEEVENPEKTPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT PSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SSLFYVASREGHLFELISMIFVKFHIPLAFIYC LFMYALSLYSPPFSTGIGFVTLTGVBAYVLFII WDKKPRWFRIMSEKITRTLQIBLEVVPEEDKL 978 2328 A 8261 2 2165 RGGSRCVLGKLLGQLLCFQSERCVVEFFGLL LADAASMSGVRAVRISIESACEKQVFBVGLG GFTYLPPLSSKPGSSMOGUPP LADAASMSGVRAVRISIESACEKQVFBFGLIGL VKBLYKKPLDFSNSQNLARLAQRIDFSOGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK PGPSLWPWDSVRNNISALTEMCVLYDVLSI VRNKKFMIDPSVQDALPPKQNFQTLQLISK KKSLAGAAQILLKGAERLTKSVTENOENKLQ RPDFNSELLRLRQHWKLRKYGDKLIGDLSYRS AGSLPPHHGTFEVKNITDLDLDKKPEDYCLI DVQIPSDLEGSAYIKVSQKQAPDIGDLGTVN LFKRPLPKSKFGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISSPFSSLG LISISCHSSNDKKSQKFATREGCCPEDHLYVLE LISISCHSSNDKKSQKFATREGCCPEDHLYVLE LISISCHSSNDKKSQKFATREGCCPEDHLYVLE LISISCHSSNDKKSQKFATREGCCPEDHLYVLE LISISCHSSNDKKSQKFATREGCCPEDHLYVLE]					
KLNNINDILKSVFLIPPIFCLGRGLIDMYKNQ AMADALERFGENFFVSPLSWDLVGRNLFAM AVEGVYFLITVLIQYRFIRPRPVNAKLSPLN DEDEDVRRENQRLDGGGQNDILEIKLTKIY RRRKRFAVDRICVGIPPGEGFGLLGVNGAGK SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV HQNMGYCPGPAITELLTGETVSPFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRLWNCALSVYKEGRSVVLTSHSMEEC EALCTTRMAIMVNGRFRCLGSVQHLKNRFGD GYTTVVRIAGSNPDLKFVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSGKKRLH EEDYSVSGTTLDQVFVVNFAKDQSDDPHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 11 EPGYSVSGTTLDQVFVVNFAKDQSDDPHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCQVISLFGALSYAELGTTIKKSGGH YTVILEVFGPLPAFVRVWVELLIRPAATAVIS LAFGRYLLEPFFPCCEPPELAKLTAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWPYLNFVTEEVENPEKTIPLACISM AIVTIGYVLTNAVFTTINAEELLISAVAVT FESERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLEFELSMHYRKHTPLPAVIV LHPLTMIMI-SGDLDSLLNSFAR WLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLVSDPFSTGIGEVITLTGVPAYYLFII PSFRLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLEFELSMHYRKHTPLPAVIV LHPLTMIMI-SGDLDSLLNSFAR WLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLVSDPFSTGIGEVITLTGVPAYYLFII PSFRLLGNFSLAVFIFVALSCFGSMNGGVFAV SRLFYVASREGHLEFELSMHYRKHTPLPAVIV LHPLTMIMI-SGDLDSLLNSFAR WLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLVSDPFSTGIGEVITLTGVCYPYYLFI PSFRLLGNFSLAVFIFVALSGRSSADQDDEGGVVK FORSLURLRLQDHVLKLKVORULJODLSYRS AGSLFPHIGTFVVKNDLJOLDKKIPEDYCPL DVGIPSDLEGSAYIKVSIQKQAPDIGLGTVN LKRRPLFKSKFGSPHWQTKLLEAAQNIVLLCKEI FAQLSREAVQUKSQVPHIVVKNQILGDLSYPFSLQ LSISLCHSSNDKKSQKFATEKGCCPEDHLYVLE LSISLCHSSNDKKSQKFATEKGCCPEDHLYVLE LSISLCHSSNDKKSQKFATEKGCCPEDHLYVLE LSISLCHSSNDKKSQKFATEKGCCPEDHLYVLE LSISLCHSSNDKKSQKFATEKGCCPEDHLYVLE LSISLCHSSNDKKSQKFATEKGCCPEDHLYVLE LSISLCHSSNDKKSQKFATEKGCCPEDHLYVLE LSISLCHSSNDKKSQKFATEKGCCPEDHLYVLE LSISLCHSSNDKKSQKFATEKGCCPEDHLYVLE LSISLCHSSNDKKSQKFATEKGCCPEDHLYVLE LSISLCHSSNDKKSQKFATEKGCCPEDHLYVLE]					
AMADALERFGENRFVSPLSWDLVGRNLFAM AVEGVVFELITVLIQYRFIRPRPVNAKLSPLN DEDEDVRRERQRILDGGGQNDILEIKELTKIY RRKRKPAVDRICVGIPPGECFGILLGVNGAGK SSTFKMLTGDITVTRGDAFLNRISILSNIHEV HQNMGYCPGPDAITELLTGREHVEFFALLRG VPEKEVGKVGEWARKLGLVKYGEK YAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTISHSMEE EALCTRNAIMVNGRFRCLGSVQHLKNRRGD GYTTVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQVQLPSSLSAHFISLGSKKRLH EDYSVSQTTLDQVFVNFAKDQSDDHLKDL SLHKNQTVVDVAVLTSFLDEKVKESYV SGNKRKLSTAMALIGGPPVVFLDEFTGMD PKARRFLUQPLSSLAHFISLGSKKRLH EDYSVSQTTLDQVFVNFAKDQSDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV TLLRGVSIIGTIIGAGFFSKGVLQNTGSVGM SLTIWTVCGVLSLFGALSVAELGTTIKKSGGH YTLQENVNGRLPSLGNKEPPQEKVQLKRKV TLLRGVSIIGTIIGAGFFSKGVLQNTGSVGM SLTIWTVCGVLSLFGALSVAELGTTIKKSGGH YTLLEVFGPLPAFPVVELLIRPAATAVIS LAFGRYLEPFFIQCEPELAIKLITAVGITVVM VLNSMSVSWSARQIFLTFCKLTALIHIVPGW MQLKGGTQNFKDAFSGRDSSITRLPLAFYY MYAYAGWFYLNFVTEEVENPEKTPLAICISM AIVTIGVYLTNVAYFTINAEELLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPELLSMHTVRKHTFLPAVIV LHPLTMIMEFSGDLDSLTNLFSFARWFIGLA VAGLIYLRYKCPDMHRFFKVPLFIPALFSFTC LFMVALSLYSDPSTGIGVTLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PKASKELSSLASKAVGNAVARISESACEKQVHEVGLD GTETYLPPLSMSQNLALQRQIDFSQGSGSEE EAAGTEGDAQGWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI KKSLAGAAQILLKAQRIDFSQGSSGSEE EEAAGTEGDAQGWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDAAPPRQNPQTLQLISK KKSLAGAAQILLKAGRELTKSVTENQENKLQ RDPNSELELRLGPWKKGOKLIGDLSYRS AGSLFPHHGTFEVIKNTDLDLKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPOSHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVYKNQISQFFPSLQ LSISLCHSSNDKKSQKFATEKQOCEDBILYQLE LSISLCHSSNDKKSQKFATEKQOCEDBILYQLE LSISLCHSSNDKKSQKFATEKQOCEDBILYQLE LSISLCHSSNDKKSQKFATEKQOCEDBILYQLE LSISLCHSSNDKKSQKFATEKQOCEDBILYQLE LSISLCHSSNDKKSQKFATEKQOCEDBILYQLE LSISLCHSSNDKKSQKFATEKQOCEDBILYQLE LSISLCHSSNDKKSQKFATEKQOCEDBILYQLE LSISLCHSSNDKKSQKFATEKQOCEDBILYQLE		1				į	
AVEGVYFLITVLJQYRFFIRPRPVNAKLSPLN DEDEDVRRERQRILDGGQNDILEIKELTKIY RRRKKPAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGDITVTRGDAFLNRRSILSNIHEV HQNMGVCPGPJAITELLTGREHVEFFALLRG VPEKEVGKVGEWARKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRRGG GYTTVVVIGSSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQVQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLIHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIPPCVPTMVRKVSTISKG YLQGNVNGRIPSLGNKEPFGQEKVQLKRKV TLIRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCQVLSLFGALSYAFLGTTIKKSGGH YTYILEVFGPLFAFVRVWVELLIRFAATAVIS LAFGRYLEPFFIQCEFELAIKLITAVGITVM VLNSMSVSMARIQIETTCKLTALIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVERPEKTIPLAICISM AUVTIGYVLTNAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMGGGVFAV SRLFYVASRGGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKGPDMHRPFKVPLIFPALFSFTC FMVALSLYSDPFSTGIGVTILTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVYPEEDKL FMVALSLYSDPFSTGIGVTILTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVYPEEDKL FRIGGSLGCKLOKLIGGLICGSERCVRFPEGIL RHRGCGLLSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISESACEKQVHEVGLD GTETT_PPLSMSQNLARQRIDFSQGSSEE EEAAGTEGDAQGWPGAGSSADQDEEGVVK FQPSLWPMDSVRNNLRSALTEMCVLYDVLSI VRDKKFMILDPVSQDALPPKQNPQTIQLISK KKSLAGAAQILLKGAERLTKSVTENOENKLQ RDPNSELLRLRQHWKLRKVGGKII.GDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQPSDLEGSAYIKVSIQVQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQUKSQVPHIVYKNQISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCGEDHLYVLE			1				1
DEDEDVRRERORILOGGGQNDILEIKELTKIY RRKRKPAVDRICVGIPPGECFGLLGVNGAGK SSTEKMLTGDTTVTRODAFLNRNSILSNIHEV HQNMGYCPGFDAITELLTGREHPEFFALLRG VPEKEVGKVGEWAIKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTTVVRIAGSNPDLKPVQDFFGLAFFGSVPK EKHENML,QVQLPSSLSALARIFSILSQSKKRLH EDYSVSQTTLDQVFVNFAKDQSDDHLKDL SLHKNQTVDVVAVLTSLDDEKVKESTY 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGDEKVCESTY TLLRGVSUIIGTIGAGFISFKGVLQNTGSVGM SLTIWTVCGVLSLFGALSVAELGTTIKKSGGH YTYLEVFGPPAPFVVEWLLIRPAATAVIS LAFGRYILEFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTALIIIVPGV MQALKGGTQNFKDAFSGRDSSTIRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AUTTIGVYLTNAVFTITABELLLSSNAVAVT FSERLLGNESLAVPFVALSCFGSMNGGVFAV SRLFYVASREAHPVALSCFGSMNGGVFAV SRLFYPASRGVANGELTITGVATA TFSELLARGHVALARGRIFTGSGSSEE EAAGTEGDAQGWPGAGSSADQDDEGGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPFRQNPQTLQLISK KSSLAGAAQILLKGARLTKSVTENQENKLQ RDPNSELLALRGCHKKKGOKLIGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSFFHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHVVKNQISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCFEDHLYVLE							
RRKRRPAVDRICVGIPPGEGEGILIGVNGAGK SSTFKMLTGDTTVTRODAFLNRNSILSNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFFGSVPK EKHRMLQYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDHLKDL SLHKNQTVVDVAVLTSFLODEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIGGGIFISPKGVLQNTGSVGM SLTWTVCGVLSIFGLNXEPPGGEKVQLKRKV TLLRGVSIIIGTIGGGIFISPKGVLQNTGSVGM SLTWTVCGVLSIFGLNXEPPGGEKVQLKRKV TLLRGVSIIIGTIGGGIFISPKGVLQNTGSVGM SLTWTVCGVLSIFEDXYSELGTIFKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEPFFQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQFLTFCKLTALIILIVPGV MQLIKGQTONFKDAFSGRDSSITRLPLAFYYG MYAYAGGWPYLNFVETVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLIGNFSLAVPIFVALSCGSMNGGVFAV SRLFYVASREGHLEFILSMIHVRKHTPLAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRILQIILEVVPEEDKL PKGCOLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESAACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNARSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTILQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKLQDLSYRE AGSLFPHHGTTEVIKNTDLDLKKPEDVCPL DVQMPSDLEGSAJVKVSIQKAAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKPATTEKQCPEDBILJYLLE FAQLSSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKPATTEKQCPEDBILJYLLE FAGLSSNDKKSQKPATTEKQCPEDBILJYLLE FAGLSSNDKKSQKPATTEKQCPEDBILJYLLE FAGLSSNDKKSQKPATTEKQCPEDBILJYLLE FAGLSSNDKKSQKPATTEKQCPEDBILJYLLE FAGLSSNDKKSQKPATTEKQCPEDBILJYLLE FAGLSSNDKKSQKPATTEKQCPEDBILJYLLE FAGLSSNDKKSQKPATTEKQCPEDBILJYLLE FAGLSSNDKKSQKPATTEKQCPEDBILJYLLE FAGLSSNDKKSQKPATTEKQCPEDBILJYLLE FAGLSSNDKKSQKPATTEKQCPEDBILJYLLE FAGLSSNDKKSQKPATTEKQCPEDBILJYLLE FAGLSSNDKKSQKPATTEKQCPEDBILLTRAPHTEVPLITTER FAGLSCHARM							3
HONMGYCPOPAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFFGSVPK EKHRMLQYQLPSSLSSLARIFSILSQSKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMYKRPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCQVLSLFGALSYAELGTTIKKSGGH YTYTLLEVFGPLAPKVWELLIRFAATAVIS LAFGRYILEPFFQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTPLAICISM AIVTIGVVLTNVAYTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLSLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFPALFSFTC LFFWALSLYSDPFSTGIGFVILTGVPAYYLFII WDKKFRWFRIMSKSKITRTLQIILEVYPEEDKL PLADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYVDVLSI VRDKKFMTILDPVSQDALPFKQNPQTLQLISK KKSLAGAQDILLKGRATKVGDKLIGDLSYRS AGSLFPHIGITEVIKNTDLDLDKKIPEDYCPL DVIPESDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPTPSLQ LISILCHSSNDKKSQKAPTIEKQCPDHILYVLE FAQLSSREAVQIKSQVPHIVVKNQIISQPTPSLQ LISILCHSSNDKKSQKAPTIEKQCPDHILYVLE	Ì		ľ				1 - 7
977 2327 A 8260 3 1567 IPGSTISFSLCEIPPCVPTMYRKPVVSTISKGG WOLKENTAMALIGATION OF THE CONTROL ON THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF TH							
977 2327 A 8260 3 1567 IPGSTISFSLCEIPPCVPTMYRKPVVSTISKGG WOLKENTAMALIGATION OF THE CONTROL ON THE CONTROL OF THE CONTROL OF THE CONTROL OF THE CONTROL OF TH							HONMGYCPOFDAITELLTGREHVEFFALLRG
SGGNKRKLSTAMALIGGPPVYFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVOHLKNRFGD GYTTVVRIAGSNPDLKPVQDFFGLAFFGSVPK EKHRNMLQYQLPSSLSSLARFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLC:FPPCVPTMYRKPVYSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISFKGVLQNTGSVGM SLTIWTVCQVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFQCEIPELAIKLITAVGITVVM VINNSVSWSARQIGIFLTFCKLTAILIIVPGV MQLIKGGTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSINAVAVT FSERLLGNFSLAVPIFVALSCFGSMNIGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSILNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFPALFSFTC LFMVALSLYSDFFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PRINGSCOLLSRLSSKLSGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLORIDFSQGSGSEE EEAAGTEGDAQEWGGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPFKQNFQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSALVKSQKFATEKQCPEDHLYVLE FAQLSREAVQIKSQVPHIVKNQIIGGPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
977 2327 A 8260 3 1567 IPGSTISFSLEFFPFCVPTMVRKPVSTISKGG YTUVINAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH EDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 978 2327 A 8260 3 I567 IPGSTISFSLEFFPFCVPTMVRRPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTITVTVCQVLSLFGALSYAELGTTIKKSGGH YTYLLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLITFCKLTALLIIIVPGV MQLIKGGTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLHVTEVEVEPPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFFPALFSFTC LFMYALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQDLCFQSERCVRFFEGLL RHRGGGLLSSRLSAGKPPLRTSFFGSWQLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAOEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPRQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTTEVIKNTDLDLDKKIPEDYCPL DVQRSDLEGGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISGPFFSLQ LSISLCHSSNDKKSQKFNATEKQCPEDHLYVLE							
977 2327 A 8260 3 1567 IPGSTIPSLSSLARIFSILSQSKKRLH EDYSVSQTTIDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTIFSLCFIFFPCVFTMVRRPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YYVILEVFGPLPAFVRVWVELLIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFTJFCKLTAILIIIVPGV MQLKGGTONFKDAFSGRDSSTIRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSSLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKFRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQISQFPFSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE		1					PKARRFLWNCALSVVKEGRSVVLTSHSMEEC
EKHRMI QYQLPSSLSSLARIFSILSQSKKRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNQTVVDVAVLTSFLQDEKVKESYV 977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIGTIGAGIFISPKGVJQNTGSVGM SLTIWTVCQVLSLFGALSYAELGTTIKKSGGH YTYLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPFLISMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFINSEKITRIQJILEVPEEDKL PWDKKPRWFINSEKITRIQJILEVPEEDKL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRSIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYNS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQPFDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYYLE							
977 2327 A 8260 3 1567 IPGSTISPSLCFIPPPCVPTWYRKPVVSTISKGG PICQNVNGRLPSLGNKEPPGQEKVLESYV 978 2328 A 8261 2 2165 RGGSLRCVLGKLGVLLSVLSVLSVLSVLSVLSVLSVLSVLSVLSVLSVLSVLS							GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK
977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQRNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLACISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGYFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPLSMSQNLARLAQRIDFSQGSGSE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPRSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	}						
977 2327 A 8260 3 1567 IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNYNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTITIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VINSMSVSWSARIQIFLTFCKLTAILIINVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGRISSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL RHGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPRSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILLEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARQIFLITFCKLTAILIINPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVASLLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAQIILKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLPFHIGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQFPFSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLTYVLE							
TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYLLEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL RHGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETTYLPPLSMSQNLARLAQRIDFSQGSGEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQFFPSLQ LSISLCHSSNDKKSQVFNTERAQCPEDHLTYVLE	977	2327	A	8260	3	1567	
SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFIQCEPBLAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIITVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PRGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEGGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTIGGFVITLTGVPAYYLFII WDKKPRWFIMSEKITRTLQIILEVVPEEDKL PROSERVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSE EAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE		İ			l		
LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFAR WLFIGLA VAGLIYLRYKCPDMIRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL PRICE CONTROLLSRILSGREPPLRTSFFGSWQLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE				ĺ			
VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFFPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							I
MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWYTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							I
MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMHVKKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKPATEKQCPEDHLYVLE	1						
AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFAR WLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
FSERLLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFAR WLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETTYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
LHPLTMIMLFSGDLDSLLNFLSFAR WLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	}				1		
VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE		1					
LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL 978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
978 2328 A 8261 2 2165 RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1]	J .		
RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	978	2328	A	8261	2	2165	
LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQIILKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							
EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQIILKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE
FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE		(i i				EEAAGTEGDAQEWPGAGSSADODDEEGVVK
VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI
KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE							VRDKKFMTLDPVSQDALPPKONPOTLOLISK
RDFNSELLRLQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1						KKSLAGAAQILLKGAERLTKSVTENQENKLQ
AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	1	ł			}		RDFNSELLRLRQHWKLRKVGDKILGDLSYRS
DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE		}					AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL
FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE		1					DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN
FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	ļ	1					LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI
LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE	}]					FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ
HNLHLLIREFHKOTLSSIMMPHPASAPFGHKR							LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE
	L	<u></u>				L	HNLHLLIREFHKQTLSSIMMPHPASAPFGHKR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion MRLSGPQAFDKNEINSLQSSEGLLEKIIKQAK HIFLRSRAAATIDSLASRIEDPQIQAHWSNIND VYESSVKVLITSQGYEQICKSIQLQLNIGVEQI
						RVVHRDGRVITLSYQEQELQDFLLSQMSQHQ VHAVQQLAKVMGWQVLSFSNHVGLGPIESIG NASAITVASPSGDYAISVRNGPESGSKIMVQF PRNQCKDLPKSDVLQDNKWSHLRGPFKEVQ WNKMEGRNFVYKMELLMSALSPCLL
979	2329	A	8289		1053	FVWNPRGGRKRRRQAAVTQAATRASGTPSP RDGTMTQGKLSVANKAPGTEGQQQVHGEKK EAPAVPSAPPSYEEATSGEGMKAGAFPPAPTA VPLHPSWAYVDPSSSSSYDNGFPTGDHELFTT FSWDDQKVRRVFVRKVYTILLIQLLVTLAVV ALFTFCDPVKDYVQANPGWYWASYAVFFAT YLTLACCSGPRRHFPWNLILLTVFTLSMAYLT GMLSSYYNTTSVLLCLGITALVCLSVTVFSFQ TKFDFTSCQGVLFVLLMTLFFSGLILAILLPFQ YVPWLHAVYAALGAGVFTLFLALDTQLLMG NRRHSLSPEEYIFGALNIYLDIIYIFTFFLQLFG TNRE
980	2330	А	8305	59	857	ASQLPDYSISPPSLPPRISFHPSPTLARVAMAEP SEATQSHSISSSSFGAEPSAPGGGSPGACPAL GTKSCSSSCAVHDLIFWRDVKKTGFVFGTTLI MLLSLAAFSVISVVSYLILALLSVTISFRIYKSV IQAVQKSEEGHPFKAYLDVDITLSSEAFHNY MNAAMVHINRALKLIIRLFLVEDLVDSLKLA VFMWLMTYVGAVFNGITLLILAELLIFSVPIV YEKYKTQIDHYVGIARDQTKSIVEKIQAKLPG IAKKKAE
981	2331	A	8308	186	1337	TRMSRHEGVSCDACLKGNFRGRRYKCLICYD YDLCASCYESGATTTRHTTDHPMQCILTRVD YDLCASCYESGATTTRHTTDHPMQCILTRVD FDLYYGGEAFSVEQPQSFTCPYCGKMGYTET SLQEHVTSEHAETSTEVICPICAALPGGDPNH VTDDFAAHLTLEHRAPRDLDESSGVRHVRR MFHPGRGLGGPRARRSNMHFTSSSTGGLSSS QSSYSPSNREAMDPIAELLSQLSGVRSAGGQ LNSSGPSASQLQQLQMQLQLERQHAQAARQ QLETARNATRRTNTSSVTTITQSTATTNIAN TESSQQTLQNSQFLLTRLNDPKMSETERQSM ESERADRSLFVQELLLSTLVREESSSDEDDR GEMADFGAMGCVDIMPLDVALENLNLKESN KGNEPPPPPL
982	2332	Α	8315	1	1004	GSTHASADAWAQWFCTEALVMGAPVWYLV AAALLVGFILFLTRSRGRAASAGQEPLHNEEL AGAGRVAQPGPLEPEEPRAGGRPRRRDLGS RLQAQRRAQRVAWAEADENEEEAVILAQEE EGVEKPAETHLSGKIGAKKLRKLEEKQARKA QREAEEAEREERKRLESQREAEWKKEEERLR LEEEQKEEEERKAREEQAQREHEEYLKLKEA FVVEEEGVGETMTEEQSQSFLTEFINYIKQSK VVLLEDLASQVGLRTQDTINRIQDLLAEGTIT GVIDDRGKFIYITPEELAAVANFIRQRGRVSIA ELAQASNSLIAWGRESPAQAPA
983	2333	A	8320	244	1420	RRRWRARGGLVPTLAWAEATGAYVPGRDKP DLPTWKRNFRSALNRKEGLRLAEDRSKDPHD PHKIYEFVNSGVGDFSQPDTSPDTNGGGSTSD TQEDILDELLGNMVLAPLPDPGPPSLAVAPEP CPQPLRSPSLDNPTPFPNLGPSENPLKRLLVPG EEWEFEVTAFYRGRQVFQQTISCPEGLRLVGS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EVGDRTLPGWPVTLPDPGMSLTDRGVMSYV RHVLSCLGGGLALWRAGQWLWAQRLGHCH TYWAVSEELLPNSGHGPDGEVPKDKEGGVF DLGPFIVGSLGPPDLIFFTEGSGRSPRYALWFC VGESWPQDQPWTKRLVMVKVVPTCLRALVE
		ı				MARVGGASSLENTVDLHISNSHPLSLTSDQY KAYLQDLVEGMDFQGPGES
984	2334	A	8321	1	1243	ANMAPVEHVVADAGAFLRHAALQDIGKNIY TIREVVTEIRDKATRRRLAVLPYELRFKEPLPE YVRLVTEFSKKTGDYPSLSATDIQVLALTYQL EAEFVGVSHLKQEPQKVKVSSSIQHPETPLHIS GFHLPYKPKPPQETEKGHSACEPENLEFSSFM FWRNPLPNIDHELQELLIDRGEDVPSEEEEEEE NGFEDRKDDSDDDGGGWITPSNIKQIQQELE QCDVPEDVRVGCLTTDFAMQNVLLQMGLHV LAVNGMLIREARSYILRCHGCFKTTSDMSRV FCSHCGNKTLKKVSVTVSDDGTLHMHFSRNP KVLNPRGLRYSLPTPKGGKYAINPHLTEDQRF PQLRLSQKARQKTNVFAPDYIAGVSPFVENDI SSRSATLQVRDSTLGAGRRRLNPNASRKKFV KKR
985	2335	Α	8322	352	529	RRNNIRQFIMKVCISGQARWLTPVVPVLWET
986	2336	A	8325	89	1172	EAGRSLELKSLRPAWATWGNPISTKINK KMNPTDIADTTLDESIYSNYYLYESIPKPCTKE GIKAFGELFLPPLYSLVFVFGLLGNSVVVLVL FKYKRLRSMTDVYLLNLAISDLLFVFSLPFWG YYAADQWVFGLGLCKMISWMYLVGFYSGIF FVMLMSIDRYLAIVHAVFSLRARTLTYGVITS LATWSVAVFASLPGFLFSTCYTERNHTYCKT KYSLNSTTWKVLSSLEINILGLVIPLGIMLFCY SMIIRTLQHCKNEKKNKAVKMIFAVVVLFLG FWTPYNIVLFLETLVELEVLQDCTFERYLDYA IQATETLAFVHCCLNPIIYFFLGEKFRKYILQL FKTCRGLFVLCQYCGLLQIYSADTPSSSYTQS TMDHDLHDAL
987	2337	A	8326	3	470	SLSAMRFLAATFLLALSTAAQAEPVQFKDC GSVDGVIKEVNVSPCPTQPCQLSKGQSYSVN VTFTSNIQSKSSKAVVHGILMGVPVPFPIPEPD GCKSGINCPIQKDKTYSYLNKLPVKSEYPSIK LVVEWQLQDDKNQSLFCWEIPVQIVSHL
988	2338	A	8335	1205	323	VIKMALAARLLPQFLHSRSLPCGAVRLRTPA VAEVRLPSATLCYFCRCRLGLGAALFPRSAR ALAASALPAQGSRWPVLSSPGLPAAFASFPAC PQRSYSTEEKPQQHQKTKMIVLGFSNPINWV RTRIKAFLIWAYFDKEFSITEFSEGAKQAFAH VSKLLSQCKFDLLEELVAKEVLHALKEKVTS LPDNHKNALAANIDEIVFTSTGDISIYYDEKG RKFVNILMCFWYLTSANIPSETLRGASVFQVK LGNQNVETKQLLSASYEFQREFTQGVKPDWT IARIEHSKLLE
989	2339	A	8349	67	185	MSGFIHQLLIQNLFCVYHTRLKTSQGLCLLSL KSLHPMS
990	2340	A	8361	210	1115	ASPFLRPQGHDSGEREPFSQTPGLMQPFSIPVQ ITLQGSRRRQGRTAFPASGKKRETDYSDGDPL DVHKRLPSSTGEDRAVMLGFAMMGFSVLMF FLLGTTILKPFMLSIQREESTCTAIHTDIMDDW LDCAFTCGVHCHGQGKYPCLQVFVNLSHPG QKALLHYNEEAVQINPKCFYTPKCHQDRNDL LNSALDIKEFFDHKNGTPFSCFYSPASQSEDVI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LIKKYDQMAIFHCLFWPSLTLLGGALIVGMV RLTQHLSLLCEKYSTVVRDEVGGKVPYIEQH OFKLCIMRRSKGRAEKS
991	2341	A	8369	9	921	SSVVEFSALSVSMACLSPSQLQKFQQDGFLVL EGFLSAEECVAMQQRIGEIVAEMDVPLHCRT EFSTQEEEQLRAQGSTDYFLSSGDKIRFFFEK GVFDEKGNFLVPPEKSINKIGHALHAHDPVFK SITHSFKVQTLARSLGLQMPVVVQSMYIFKQP HFGGEVSPHQDASFLYTEPLGRVLGVWIAVE DATLENGCLWFIPGSHTSGVSRRMVRAPVGS APGTSFLGSEPARDNSLFVPTPVQRGALVLIH GEVVHKSKQNLSDRSRQAYTFHLMEASGTT WSPENWLQPTAELPFPQLYT
992	2342	A	8370	906	4	MALSGNCSRYYPREQGSAVPNSFPEVVELNV GGQVYFTRHSTLISIPHSLLWKMFSPKRDTAN DLAKDSKGRFFIDRDGFLFRYILDYLRDRQVV LPDHFPEKGRLKREAEYFQLPDLVKLLTPDEI KQSPDEFCHSDFEDASQGSDTRICPPSSLLPAD RKWGFITVGYRGSCTLGREGQADAKFRRVPR ILVCGRISLAKEVFGETLNESRDPDRAPERYTS RFYLKFKHLMGAPASNFILGFWGLGQNQDK HPVNIYLQQRSVIRPDLTSKKAGDLKGKGDA QEVSRRRWLGDPEHL
993	2343	A	8379		2794	MRMQRHKNDTMDFGDSGKFIGGGVLCLLHQ SNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIE DMVTTASTYLFEATEKRFFFKNVSILIPENWK ENPQYKRPKHENHKHADVIVAPPTLPGRDEP YTKQFTECGEKGEYIHFTPDLLLGKKQNEYG PPGKLFVHEWAHLRWGVFDEYNEDQPFYRA KSKKIEATRCSAGISGRNRVYKCQGGSCLSRA CRIDSTTKLYGKDCQFFPDKVQTEKASIMFM QSIDSVVEFCNEKTHNQEAPSLQNIKCNFRST WEVISNSEDFKNTIPMVTPPPPPVFSLLKIRQRI VCLVLDKSGSMGGKDRLNRMNQAAKHFLLQ TVENGSWVGMVHFDSTATIVNKLIQIKSSDER NTLMAGLPTYPLGGTSICSGIKYAFQVIGELH SQLDGSEVLLLTDGEDNTASSCIDEVKQSGAI VHFIALGRAADEAVIEMSKITGGSHFYVSDEA QNNGLIDAFGALTSGNTDLSQKSLQLESKGLT LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPP SISLWDPSGTIMENFTVDATSKMAYLSIPGTA KVGTWAYNLQAKANPETLTITVTSRAANSSV PPITVNAKMNKDVNSFPSPMIVYAEILQGYVP VLGANVTAFIESQNGHTEVLELLDNGAGADS FKNDGVYSRYFTAYTENGRYSLKVRAHGGA NTARLKLRPPLNRAAYIPGWVVNGEIEANPP RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSL PLPDQYPPSQITDLDATVHEDKIILTWTAPGD NFDVGKVQRYIIRISASILDLRDSFDDALQVN TTDLSPKEANSKESFAFKPENISEENATHIFIAI KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPT PTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIV
994	2344	A	8385	231	644	INSSPRTGRDHQELNLHTERDSRSQRAVLKIP RQNPGIFYWIFLPSRSHSASHGSRQRQVSCQG TQDEILKMRNTFAELKNSLEALSSRMDQAEE RIGTQAGVQWRDHGSLQPQPPEFKQCFHLSL
995	2345	A	8390	194	3421	PSSWDYRACLS AWRKSSVVPPRGTRRGEKSDQDKSGQKNKR

nucle- cetide ce	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
eotide sequence	NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine
uence 914 ng to first amino acid residue of peptide residue of peptide sequence T-Threonine, V-Valine, W-Trytophan, Y-Stop codon, Possible nucleotide deletion, \=\text{possible} peptide sequence peptide sequence peptide sequence peptide sequence pertia p			!				
amino acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence per sequence peptide sequence per seq	seq-	uence		09/496	correspondi		
residue of peptide sequence Y=Tyrosine, X=Luknown, *=Siop codon,	uence			914			
peptide sequence Possible nucleotide deletion, "possible nucleotide insertion							
sequence nucleotide insertion						sequence	
DPLSMKQSPALAPEERCRAGSPKPVLRAI NNMGNGCSQKLATANLEPILLVLIPCICA LLIELLSVYGTLQKYYFKSNGSEPLVTDG QGSDVILTNTTYNQSTVYSTAHPDQHVPAV TDASLPGDQSHRNTSACMNTHSQCQMLP ATLTPLSVVRMEMEKFLKFFTYLHRLSK QHIMLFGCTLAFPECIIDGDDSHGLLPCRSF AAKEGCESVLGMVNYSWPDFLRCSQFRNG ESSNVSRICFSPQENGKQLLGGGENFLC GICPGKLQCNGYMDCDDWSDEAHCNCSE FHCHTGKCLNYSLVCDGYDDCGDLSDEQN DCNPTTEHRCGDGRCLAMEWVCDGHIDC KSDEVNCSCHSQGLVCERNGQCIPSTFQCT DEDCKDGSBLEDRNSINNCSQCEPITLELO KSDEVNCSCHSQGLVCERNGQCIPSTFQCT DEDCKDGSSLCDPNNSINNCSQCEPITLELO NLPYNSTSYPNYFGHRTQKEASISWESSLF LVQTNCYKYLMFFSCTILLVPKCDVTICEH CRALCEHSKERCESVLGIVGLQWPEDTDC FPEENSDNQTCLMPDEYVEECSPSHFKCRS CVLASRGCDGQADCTUSINNVNSSSFLMVHRAATE VCADGWCELSQLACKQMGLGFSVTKLL QEKEPRALTHISNWESLNGTTHELLVNC CESRSKISLLCTKQDCGRPAARMNKRILG TSRFGRWPWQCSLQSEPSGHIGCSVLIAKK VLTVAHCFGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEL DISSTGVRPVCLPRGVLEPDTYCYTTG GHMGNKMPFKLQEGEVRIJSLEHCQSYFDI TITITMICAGYBEGTVDSCMGDSGGPLVC PGGRWTLFGLTSWGSKGPLVC PGGRWTLFGLTSWGSKGPLVC PGGRWTLFGLTSWGSKGPLVC PGGRWTLFGLTSWGSKGPLVC PGGRWTLFGLTSWGSKGPLVC PGGRWTLFGLTSWGSKGPLVC PGGRWTLFGLTSWGSKGPLVC PGGRWTLFGLTSWGSKGPLVC PGGRWTLFGLTSWGSKGPLVC PGGRWTLFGLTSWGSKFRSRSSRSRSSRSSRSSRSSRSSRSSRSSRSSRSSRSSRS	})				
NNMGNGCQKLATANLLRLLLVLPCICA LLLEILLSYVGTLQKVYFKSNGSEPLVTDG QGSDVLINTIYNQSTVVSTAHPDQHVPAV TDASLPGDQSHRNTSACMINTHSQCQMLP ATLTPLLSVVRNMEMEKFLKFFTYLHRLSK QHMLFGCTLAFPECIIDGDDSHGLLPCRSF AAKEGCESVLGMVNYSWPDFLRCSQFRING ESSNVSRICFSPQQENGKQLLCGRGENFLC GICPGKLQCNGYNDCDDWSDEAHCNCSE FHCHTGKCLNYSLVCDGYDDCGDLSDEQ DCNPTTEHRCGDGRCLAMEWVCDGDHDC KSDEVNCSCHSQGLVECRNGQCIPSTFQCL DEDCKDGSDEENGSVQTGSCQEGDQRCLY CLDSCGGSSLCDPINSLNNCSQCEPITLEL NLPYNSTSYPNYFGHRTQKEASISWESSLF LVQTNCYKYLMFFSCTILVPKCDVNTGEH CRALCEHSKERCESVLGIVGLQWPEDTDC: PFEENSDNQTCLMPBEVVECGSPSHCKS CVLASRRCDGQADCDDSDEENCGCKERI WECPSNKQCLKHTVICDGFDCPDYNDGE CSFCQDELLECANHACVSRDLWCDGEAD DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQEILSQLACKQMGLGEPSVTKLI QEKEPRWLTLHSNWESLIGHTLHELLVNC CESRKISLLCTKQDCGRPAARMKRILL TSPGGRWPWQCSLQSEPSGHIGGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHFRYSRAVVDYDISIVCU DISETGYVRPVCLIPPEGWLEPDTYCYTITG GHMGNKMPFKLQGEFVRISLEHICQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGFGVYSN YFVEWIKRQIYQFTLLM SVLSSEMSKTNKSKSGSRSSRSRSASSRSS FSKSRSRSLSRSRKRRLSSRSRSSYSPA RENNHPRVYQNRDFRGHNRGYRRPYYFAR NGGFYPWQQYNRGGYGNYRSNWQNYRQ SPRGGRSRSSSRSSNSRSRSRSRSSSSRSSSSRSSSSRSS					Sequence		
LLLEILLSYVGTLQKVYFKSNGSEPLVTDG QGSDVILTNTIYNQSTVVSTAHPDQHVPAV TDASLPGDQSHRNTSACMNITHSQCQMLP ATLTPLLSVVRNMEMEKFIFTYLIRLS QHIMLFGCTLAFPECIIDGDSHGLLPCRSF AAKEGCESVLGMVNYSWPDFLRCSQFRNG ESSNYSRICFSPQQENGKQLLGRGENFLC GICIPGKLQCNGYNDCDDWSDEAHCNCSE FHCHTGKCLNYSLVCDGYDDCGDLSDEQ DCNPTTHERGDGRCIAMEWVCDGDHDC KSDEVNCSCHSQGLVECRNGQCIPSTFQCT DEDCKDGSDEENCSVIQTSCQEGDQRCLY CLDSCGGSSLCDPNNSLNNCSQCEPITLELG NLPYNSTSYPNYFGHRTQKEASISWESSLF LVQTNCYKYLMFFSCTILLYRCDVNTGEH CRALCFHSKERCESVLGIVGLQWPEDTDC: FPEENSDNQTCLMPDEYVEECSPSHFKCRS CVLASRRCDQADCDDDSDEENCGCKERI WECPSNKQCLKHTVICDGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGADD DSSDEWDCVTLSINNNSSSFLWHRAATEI VCADGWQELSQLACKQMGLGEPSVTKLI QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEFSGHIGGCVLIAKK VLTVAHCFEGRENAAVWKVLGINNLDH VPMQTRFVKTILHFRYSRAVVDYDISIVEI DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQGEEVRIJSLEHCQSYFD TITITRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN VFVEWIKRQIYQTFLLN SERSHSLSTSRSKRRLSSRSRSSSSSSSSSSSSSSSSSSSSSSSS							
QGSDVILITITYNQSTVYSTAFPDQHYPAW TDASLPGDQSHRNTSACMNITHSQCQMLP ATLTPLLSVVRNMEMEKFLKFFTYLHRLS(QHMM_FGCTLAFPECIIIGDDSHGLLPCRSS AAKEGCESVLGMVNYSWPDFLRCSQFRN ESSNVSRICFSPQQENGKQLLCGRGENFLC GICIPGKLQCNGYNDCDDWSDEAHCNCSE FHCHTGKCLNYSLOGYDDCGGLSDEQ) DCNPTTEHRCGDGRCIAMEWVCDGDHDC KSDEVNCSCHSQGLVECRNQQCIPSTFQCL DEDCKDGSDEENCSVQTSCQEGDQRCLY; CLDSCGGSSLCPNNSLNNCSQCEPITLELL NLPYNSTSYPNYFGHRTQKEASISWESSLF LVQTNCYKYLMFFSCTILVPKCDVNTGEH CRALCEHSKERCESVLGIVGLQWPEDTDC: FPEENSDNQTCLMPEVVECSSPSHFKCRS CVLASRRCDGQADCDDDSDEENCGCKER; WECPSNKQCLKHTVICDGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGEAD DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQELISQLKCQMGLGGFSVTKLI QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHIGGCVLLAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTILHPRYSRAVVDYDISIVEI DISETGYVRPVCLPPPEQWLEPDTYCYTIG GHMGNKMPFKLQGESVRIJSLEHCQSYFO TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 596 2346 A 8392 199 3085 KVLSSEMSKTNKSKSGGSRSSRSRSASRSRS FSKSRSRSSLSRSRKRRLSSRSRSSSSSSSRSRSASRSRS FSKSRSRSSLSRSRKRRLSSRSRSSSSSSSRSNSNSNSNSNSNSNSNSNSNNONYRQ SPRGGRSRSSSRSRSNSRNSNONYRQ SPRGGRSRSSSRSRSNSRNSNNONYRQ SPRGGRSSSSSSSNSNSNNONYRQ SPRGGRSRSSSRSSNSNSNNONYRQ SPRGGRSSSSSSSNSNSNNONYRSS DRSRRSSSSSSSSNSNSNSNNONYRSSSPSLSRSNNONYRSS SPALKSPLQSVVVRRSSPPPSSPVKESPPSPS SPALKSPLQSVVVRRSPPPSPSPVKESPPSPS SPALKSPLQSVVVRRSPPPSSPVKESPPSPS							
ATLTPLLSVYRNMEMEKFLKFFTYLHRLSK QHIMLFGCTLAFPECIIDGDDSHGLLPCRSF AAKEGCESVLGMYNYSWPDFLRCSQFRIK ESSNVSRICFSPQENGKQLLCGRGENFLC GICIPGKLQCNGYNDCDDWSDEAHCNCSE FHCHTGKCLNYSLVCDGYDDCGDLSDEGG DCNFTTEHRCGDGRCIAMEWVCDGDHDC KSDEVNCSCHSQGLVECRNGQCIPSTFQCT DEDCKDGSDEENCSVIQTSCQEGDQRCLY/ CLDSCGGSSLCDPNNSLNNCSQCEPITLELG NLPYNSTSYPNYFGHRTQKEASISWESSLFI LVQTINCYKYLMFFSCTIL.VPKCDVNTGEH CRALCEHSKERCESVLGIVGLQWPEDTDC: FPEENSDNQTCLMPDEYVECSPSHFKCRS CVLASRRCDGQADCDDDSDEENCGCKERI WECPSNKQCLKHTVICOGFPDCPDYMDEK CSFCQDDELECAHACVSRDLWCDGEAD DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQELSQLACKQMGLGEPSVTKLI QEKEPRWLTHAWESLNGTTHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGNNLDH VFMQTRFVKTILHPRYSRAVVDYDISIVEL DISETGYVRPVCLPNPEGWLEPDTYCYTTG GHMGNKMPFKLQEGEVRISLEHCQSYFOI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYJOTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSSSSPA RERNIPRVYQNRDFRGHNG VRRPYYFRK NRGFYPWQQYNRGGYGNYRSNWQNYRQ SPRGRSRSSRSLSRSKRRSSRSRSSSSNSSNSSNSSNSSNSSNSSNSSNSSNSSNS							QGSDVILTNTIYNQSTVVSTAHPDQHVPAWT
QHIML.FGCTL.AFPECIIDGDDSHGLLPCRSF AAKEGCESVLGMVNYSWPDFLRCSQFRN ESSNVSRICFSPQQENGKQLLCGRGENFLC GICPGKLQCNGYNDCDDWSDEAHCNCSE FHCHTGKCLNYSLVCDGYDDCGDLSDEQI DCNPTTEHRCGDGRCLAMEWVCDGDHDC KSDEVNCSCHSQGLVCERNGQCIPSTFQCC DEDCKDGSDEENCSVIQTSCQEGDQRCLY CLDSCGGSSLCPNNSLNNCSQCEPITLELC NLPYNSTSYPNYFGHRTQKEASISWESSLF LVQTNCYKYLMFFSCTILVPKCDVNTGEH CRALCEHSKERCESVLGIVGLQWPEDTDC: FPEENSDNQTCLMPDETYVECSPSHIFKCRS CVLASRRCDGADCDDDSDENCGKER WECPSNKQCLKHTVICDGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGEAD DSSDEWDCVTLSINVNSSSFLMYHRAATEI VCADGWQEILSQLACKQMGLGEPSVTKLI QEKEPRWLTLHSNWESLNGTTLHELLVNG CESSRKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFFMQTRFVKTIILHPRYSRAVVDVDISIVEI DISETGYVRPVCLPWFEQWLEPDTYCYTTG GHMGNKMPKLQEGEVRIISLEHCQSYFDI TITTRMCAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYQTFLIN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSSRSSRSSPSPRSRSSRSSSSPSRRSRSSSSSS							TDASLPGDQSHRNTSACMNITHSQCQMLPYH
AAKEGCESVLGMVNYSWPDFLRCSQFRMC ESSNVSRICFSPQQENGKQLLCGRGENFLC GICPGKLQCNGYNDCDDWSDEAHCNCSE FHCHTGKCLNYSLVCDGYDDCGDLSDEQN DCNPTTEHRCGDGRCIAMEWVCDGHDC KSDEVNCSCHSQGLVECRNGQCPSTFQCT DEDCKDGSDEENCSVIQTSCOEGDQRCLYT CLDSCGGSSLCDPNNSLNNCSQCEPTTLELC NLPYNSTSYPNYFGHRTQKEASISWESSLF LVQTNCYKYLMFFSCTILVPKCDVNTGEH CRALCEHSKERCESVLGIVGLQWPEDTDC: FPEENSNQTCLMPDEYVECSPSHFKCRS CVLASRCDGQADCDDDSDEENCGCKERI WECPSNKQCLKHTVICDGFPDCPDYMDEX CSFCQDDELECANHACVSRDLWCDGEAD DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQEILSQLACKQMGLGEPSVTKLI QEKEPRWLTLHSNWESLNGTTLHELLVNG CESSKISLLCTKQDCGRRPAARMKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEI DISETGYVRPVCLPNPEQWLEPDTYCYTTG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TTTTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYQTTELN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSSRSRSSRSRSRSRSRSRSRSRSRSRSR							ATLTPLLSVVRNMEMEKFLKFFTYLHRLSCY
ESSNYSICESPQQENGKQLLCGRGENFLC GICIPGKLQCNGYNDCDDWSDEAHCNCSE FHCHTGKCLNYSLVCDGYDDCGDLSDEQN DCNPTTEIRCGDGRCIAMEWCDGDHDC KSDEVNCSCHSQGLVECKNGQCIPSTFQCL DEDCKDGSDEENCSVIQTSCQEGDQRCLY CLDSCGGSSLCDPNNSLNNCSQCEPITLELG NLPYNSTSYPNYFGHRTQKEASISWESSLF LVQTNCYKYLMFFSCTILLYRCDVNTGEH CRALCEHSKERCESVLGIVGLQWPEDTDC: FPEENSDNQTCLMPDETYVECSPSHIFKCRS CVLASRRCDGQADCDDDSDEENCGCKERI WECPSNKQCLKHTVICDGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGEAD DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQEILSQLACKQMGLGEPSVTKLI QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPYSRAVVDYDISIVEL DISETGYVRPVCLPNPEQWLEPDTYCYTTG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TTTTRMLGAYESGTVDSCMGDSGGFLVCI PGGRWTLFGLTSWGSVCPSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSSRSASRSSS FSKSRSRSSSSSSNHSRVESSKRSSRSSSSYSPA RERNHPRVYQNRDFRGHNRGYRRPYYFFR NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSRSSRSHSRNSDKS DRSRRSSSSSSSNHSRVESSKRKSAKEKK KDSRPSQAAGDNQGDEVKEQTFSGGTSQD ASESSKPWPDATYGTGSASRASAVSLSSP							
GICIPGKLQCNGYNDCDDWSDEAHCNCSE FHCHTGKCLNYSLVCDGYDDGDLSDEQ DCNPTTEHRCGDGRCIAMEWCDGDHDC KSDEVNCSCHSQGLVECRNGQCIPSTFQCI DEDCKDGSDEENCSVIQTSCQGDQRCLY CLDSCGGSSLCDPNNSLNNCSQCEPITLELG NLPYNSTSYPNYFGHRTQKEASISWESSLF LVQTNCYKYLMFFSCTILVPKCDVNTGEH CRALCEHSKERCESVLGIVGLQWPEDTDC: FPEENSDNQTCLMPDEYVEECSPSHFKCRS CVLASRRCDGQACDDDSDEENCGCKERI WECPSNKQCKHTVICDGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGEAD DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQELISQLACKQMGLGEPSVYKLI QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG. TSRPGRWPQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEI DISETGYVPCLFNPEQWLEPDTYCYTTG GHMGNKMPFKLQGEVRISLEHCQSYFDI TITITMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSSSSRSRSASRSRS FSKSRSRSSLSRSRKRRLSSRSRSRSYSPA RERNHPRVYQNRDFRGHNRGYRRPYYFRR NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGSRSSRSSRSSRSSRSSNSNSSKS DRSRSSSSSSSSNSNSKVESSKRKSAKEKK KDSRPSQAAGDNQGDEVKEQTFSGGTSQD ASESSKPWPDATYGTGSASRASVSELSPR SPALKSPLQSVVVVRRSPRPSSPVPKPSPPLS			'				
FHCHTGKCLNYSLVCDGYDDCGDLSDEQY DCNPTTEHRCGDGRCIAMEWVCDGDHDC KSDEVNCSCHSQGLVECRNQCUEPSTFQCE DEDCKDGSDEENCSVIQTSCQEGDQRCLYY CLDSCGGSSLCDPNNSLNNCSQCEPITLELL NLPYNSTSYPNYFGHRTQKEASISWESSLF LVQTNCYKYLMFFSCTILVPKCDVNTIGEH CRALCEHSKERCESVLGIVGLQWPEDTDC: PPEENSDNQTCLMPDEYVEECSPSHFKCRS CVLASRRCDGQADCDDDSDEENCGCKERI WECPSNKQCLKHTVICDGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGEAD DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQEILSQLACKQMGLGEPSVTKLI QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLCTKQDCGRPAARNNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCPEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEI DISETGYVRPVCLPNPEQWLEPDTYCYTIG GHMGNKMPFKLQGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNSKSGSRSSRSSRSSRSSRSSPSARSRS FSKSRSRSLSRSRKRRLSSRSRSRSYSPAI RERNHPRVQQNRGGYONYRSNWQNYRQ SPRGRSRSRSSRSSRSSRSSRSSNSNSKS DRSRRSSSRSSSNHSRVESSKRKSAKEKK KDSRPSQAAGDNQGDEVKEQTFSGGTSQL ASESSKPWPDATYGTGSASRASVSELSPR]	•				l
DCNPTTEHRCGDGRCIAMEWVCDGDHDC KSDEVNCSCHSQGLVECRNGQCIPSTFQCT DEDCKDGSDEENCSVQTSCQEGDQRCLY. CLDSCGGSSLCDPNNSLNNCSQCEPITLELL NLPYNSTSYPNYFGHRTQKEASISWESSLF LVQTNCYKYLMFFSCTILVPKCDVNTGEH CRALCEHSKERCESVLGIVGLQWPEDTDC: FPEENSDNQTCLMPDEYVEECSPSHFKCRS CVLASRRCDGQADCDDDSDEENCGCKERI WECPSNKQCLKHTVICOGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGEAD DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWGELSQLACKQMGLGEPSVTKLI QEEEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEIL DISETGYVRPVCLPNFEQWLEPDTYCYITG GHMGNKMPFKLGECVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLIN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSRSSRSRS FSKSRSRSLSRSRKRRLSSRSRSRSYSPA RERNHPRVYQNRDFRGHNRGYRRPYFRG NRGFYPWGQYNRGGYGNYRSNWQNYQ SPRGRSRSRSPKRRSSSRSRSSRSRSSSSSRSSSSSSSSSS	1						•
DEDCKDGSDEENCSVIOTSCQEGDQRCLY CLDSCGGSSLCDPNNSLNNCSQCEPITLELL NLPYNSTSYPNYFGHRTQKEASISWESSLF LVQTNCYKYLMFFSCTILVPKCDVNTGEH CRALCEHSKERCESVLGIVGLQWPEDTDC: FPEENSDNQTCLMPDEYVEECSPSHFKCRS CVLASRRCDGQADCDDDSDEENCGCKERI WECPSNKQCLKHTVICDGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGEADI DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQELSQLACKQMGLGEPSVTKLI QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEI DISETGYVRPVCLPNPEQWLEPDTYCYTTG GHMGNKMPFKLQEGEVRIJSLEHCQSYFDI TITTRMCAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRS FSKSRSRSSLSRSRKRRLSSRSRSRSYSPA RERNHPRVYQNRDFRGHNRGYRRPYYFRR NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGSRSSRSRKRRSSSSRSHSRNSDKS DRSRRSSSSRSSNHSRVESSKRKSAKEKK KDSRPSQAAGDNQCDEVKEQTFSGGTSQL ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVFKPSPPLS							DCNPTTEHRCGDGRCIAMEWVCDGDHDCVD
CLDSCGGSSLCDPNNSLNNCSQCEPITLELG NLPYNSTSYPNYFGHRTQKEASISWESSEF LVQTNCYKYLMFFSCTIILVPKCDVNTGEH LVQTNCYKYLMFFSCTIILVPKCDVNTGEH CRALCEHSKERCESVLGIVGLQWPEDTDC: FPEENSDNQTCLMPDEYVEECSPSHFKCRS CVLASRRCDGQADCDDDSDEENCGCKERI WECPSNKQCLKHTVICDGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGEADI DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQEILSQLACKQMGLGEPSVTKLII QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEL DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGYYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSRSRSRSSPAR RERNHPRVYQNRDFRGHNRGYRRYYFRE NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGSRSSRSRSRSRSRSRSRSRSRSRSRSSRSRSRSRS							KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG
NLPYNSTSYPNYFGHRTQKEASISWESSLF, LVQTNCYXYLMFFSCTILVPKCDVNTGEH CRALCEHSKERCESVLGIVGLQWPEDTDC: PPEENSDNQTCLMPDEYVEECSPSHFKCRS CVLASRRCDGQADCDDDSDEENCGCKERI WECPSNKQCLKHTV1CDGFPDCPD YMDEK CSFCQDDELECANHACVSRDLWCDGEAD DSSDEWDCVTLSINVNSSSLMVHRAATEI VCADGWQEILSQLACKQMGLGEPSVTKLI QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEL DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TTTTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSASRSRS FSKSRSRSLSRSKKRLSSRSRSRSSSSPAR RERNHPRVYQNRDFRGHNRGYRPYYFRE NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSSRSSNHSRVESSKRKSAKEKK KDSRPSQAAGDNQGDEVKEQTFSGGTSQL ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVVRRRSPPSPVPKPSPPLS						•	DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP
LVQTNCYKYLMFFSCTILVPKCDVNTGEH CRALCEHSKERCESVLGIVGLQWPEDTDCS FPEENSDNQTCLMPDEYVEECSPSHFKCE CVLASRRCDGQADCDDDSDEENCGCKERI WECPSNKQCLKHTVICDGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGEADI DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQEILSQLACKQMGLGEPSVYTKLI QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLLAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEL DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRS FSKSRSRSLSRSRKRRLSSRSRSRSYSPA RERNHPRVQNRDFRGHNRGYRRPYYFR NRGFYPWQQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSSHSRNSDKS DRSRRSSSRSSNHSRVESSKRKSAKEKI KDSRPSQAAGDNQGDEVKEQITFSGGTSQL ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVVRRSRPPSPVFKSSPPLS							CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM
CRALCEHSKERCESVLGIVGLQWPEDTDC: FPEENSDNQTCLMPDEYVEECSPSHFKCRS CVLASRRCDGQADCDDDSDEENCGCKERI WECPSNKQCLKHTVICDGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGEAD DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWEILSQLACKQMGLGEPSVTKLI QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVTLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEI DISETGYVRPVCLPNPEQWLEPDTYCYTTG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRS FSKSRSRSLSRSRKRRLSSRSRSPSPAL RERNHPRVYQNRDFRGHNRGYRRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGSRSSSPKRRSPSPRSRSHSRNSDKS DRSRRSSSSRSSSNHSRVESSKRKSAKEKK KDSRPSQAAGDNQGDEVKEQTFSGGTSQL ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSFRPSPVPKPSPPLS	1		1				
FPEENSDNQTCLMPDEYVEECSPSHFKCRS CVLASRCDGQADCDDDSDEENCGCKERI WECPSNKQCLKHTVICDGFPDCPMDEK CSFCQDDELECANHACVSRDLWCDGEADI DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQEILSQLACKQMGLGEPSVTKLI QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMIKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCPEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEI DISETGYVRPVCLPNPEQWLEPDTYCYTTG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRS FSKSRSRSISSRSRKRRLSSRSRSRSYSPAA RERNIPPRVYQNRDFRGHNRGYRRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSSHSRNSDKS DRSRRSSSSRSSSNHSRVESSKRKSAKEKK KDSRPSQAAGDNQGDEVKEQTFSGGTSQL ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRSPRPSPVKPSPPLS		'	'				
CVLASRRCDGQADCDDDSDEENCGCKERI WECPSNKQCLKHTVICDGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGEADD DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQEILSQLACKQMGLGEPSVTKLII QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEI DISETGYVRPVCLPPPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSYSPAI RERNHPRVYQNRDFRGHNRGYRRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSSRSSNHSRVESSKRKSAKEKK, KDSRPSQAAGDNQGDEVKEQTFSGGTSQL ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRSPRPSPVKPSPPLS							FPFENSDNOTCI MPDEVVEECSPSHEKCDSCO
WECPSNKQCLKHTVICDGFPDCPDYMDEK CSFCQDDELECANHACVSRDLWCDGEAD DSSDEWDCVTLSINVNSSSFLMYHRAATEI VCADGWQEILSQLACKQMGLGEPSVTKLI QEKEPRWLITLHSNWESLINGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEL DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRS FSKSRSRSSLSRSRKRRLSSRSRSRSSSPAAR RERNHPRVYQNRDFRGHNRGYRRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSSRNHSRVESSKRKSAKEKK DRSRRSSSSRSSSNHSRVESSKRKSAKEKK KDSRPSQAAGDNQGDEVKEQTFSGGTSQL ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS							
CSFCQDDELECANHACVSRDLWCDGEADO DSSDEWDCVTLSINVNSSSFLMVHRAATEI VCADGWQEILSQLACKQMGLGEPSVTKLII QEKEPRWTLHSNWESLNGTTLHELLVNO CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEL DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSSSSSSSSSSSSSS							
VCADGWQEILSQLACKQMGLGEPSVTKLII QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDHI VFMQTRFVKTIILHPRYSRAVVDYDISIVEI DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRS FSKSRSRSLSRSRKRRLSSRSRSRSPSA RERNHPRVYQNRDFRGHNRGYRRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSRSSSNHSRVESSKRKSAKEKKS KDSRPSQAAGDNQGDEVKEQTFSGGTSQL ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS							CSFCQDDELECANHACVSRDLWCDGEADCS
QEKEPRWLTLHSNWESLNGTTLHELLVNG CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEL DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSSRSRS FSKSRSRSRSLSRSRKRLSSRSRSRSPAL RERNIPRVYQNRDFRGHNRGYRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSSRSSSNHSRVESSKRKSAKEKKS KDSRPSQAAGDNQGDEVKEQTFSGGTSQU ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS							DSSDEWDCVTLSINVNSSSFLMVHRAATEHH
CESRSKISLLCTKQDCGRRPAARMNKRILG TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEL DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSSRSRS FSKSRSRSSLSRSRKRRLSSRSRSSPSA RERNIPRVYQNRDFRGHNRGYRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSSRSSSNHSRVESSKRKSAKEKKS KDSRPSQAAGDNQGDEVKEQTFSGGTSQL ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS							VCADGWQEILSQLACKQMGLGEPSVTKLIQE
TSRPGRWPWQCSLQSEPSGHICGCVLIAKK VLTVAHCFEGRENAAVWKVVLGINNLDHI VFMQTRFVKTIILHPRYSRAVVDYDISIVEL DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSSRSSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSSPAI RERNIHPRVYQNRDFRGHNRGYRRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSSRSSSNHSRVESSKRKSAKEKS KDSRPSQAAGDNQGDEVKEQTFSGGTSQL ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS		į	1				
VLTVAHCFEGRENAAVWKVVLGINNLDH VFMQTRFVKTIILHPRYSRAVVDYDISIVEL DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSSRSSRS FSKSRSRSRSLSRSRKRRLSSRSRSSPSAI RERNHPRVYQNRDFRGHNRGYRRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSSRSSSNHSRVESSKRKSAKEKS KDSRPSQAAGDNQGDEVKEQTFSGGTSQL ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS					·		· · · · · · · · · · · · · · · · · · ·
VFMQTRFVKTIILHPRYSRAVVDYDISIVEL DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCH PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSSRSSRSSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSSPAI RERNHPRVYQNRDFRGHNRGYRRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSSRSSNHSRVESSKRKSAKEKS KDSRPSQAAGDNQGDEVKEQTFSGGTSQD ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS		ļ					
DISETGYVRPVCLPNPEQWLEPDTYCYITG GHMGNKMPFKLQEGEVRIISLEHCQSYFDI TITTRMICAGYESGTVDSCMGDSGGPLVCH PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSSRSSRSSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSYSPAI RERNHPRVYQNRDFRGHNRGYRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSSRSSSNHSRVESSKRKSAKEKS KDSRPSQAAGDNQGDEVKEQTFSGGTSQG ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS		İ	}				
TITTRMICAGYESGTVDSCMGDSGGPLVCI PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSSRSRSS FSKSRSRSRSLSRSRKRRLSSRSRSRSPSPA RERNHPRVYQNRDFRGHNRGYRRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSSRSSSNHSRVESSKRKSAKEKS KDSRPSQAAGDNQGDEVKEQTFSGGTSQG ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS							DISETGYVRPVCLPNPEQWLEPDTYCYITGW
PGGRWTLFGLTSWGSVCFSKVLGPGVYSN YFVEWIKRQIYIQTFLLN 996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSYSPAI RERNHPRVYQNRDFRGHNRGYRRPYYFRO NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSRSSRNSRVESSKRKSAKEKS KDSRPSQAAGDNQGDEVKEQTFSGGTSQD ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS							GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK
996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSYSPAL RERNHPRVYQNRDFRGHNRGYRRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSSRSSNHSRVESSKRKSAKEKS KDSRPSQAAGDNQGDEVKEQTFSGGTSQD ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS							TITTRMICAGYESGTVDSCMGDSGGPLVCEK
996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSASRSRS FSKSRSRSRSLSRSRKRRLSSRSRSRSYSPAL RERNHPRVYQNRDFRGHNRGYRRPYYFRG NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSRSSSNHSRVESSKRKSAKEKK KDSRPSQAAGDNQGDEVKEQTFSGGTSQD ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS							
FSKSRSRSILSRSRKRRLSSRSRSYSPAL RERNHPRVYQNRDFRGHNRGYRRPYYFRO NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSRSSSNHSRVESSKRKSAKEKK; KDSRPSQAAGDNQGDEVKEQTFSGGTSQD ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS	006	2246	Λ	9202	100	2005	
RERNHPRVYQNRDFRGHNRGYRRPYYFRO NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSRSSSNHSRVESSKRKSAKEKK KDSRPSQAAGDNQGDEVKEQTFSGGTSQD ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS	330	2346	A	8392	199	3083	!
NRGFYPWGQYNRGGYGNYRSNWQNYRQ SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSRSSSNHSRVESSKRKSAKEKK KDSRPSQAAGDNQGDEVKEQTFSGGTSQD ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS							
SPRRGRSRSPKRRSPSPRSRSHSRNSDKS DRSRRSSSSRSSSNHSRVESSKRKSAKEKKS KDSRPSQAAGDNQGDEVKEQTFSGGTSQD ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS							
KDSRPSQAAGDNQGDEVKEQTFSGGTSQD ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS]						SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS
ASESSKPWPDATYGTGSASRASAVSELSPR SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS							DRSRRSSSSRSSSNHSRVESSKRKSAKEKKSSS
SPALKSPLQSVVVRRRSPRPSPVPKPSPPLS							KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK
SOMEGIT PREACVOSETHOCOTDUCSES							
							SPALKSPLQSVVVRRRSPRPSPVPRPSPPLSST SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP
SKKSPVGKSPPSTGSTYGSSOKFESA A SGG							SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA
							YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE
							KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG
SQSPKRYKLRDDFEKKMADFHKEEMDDQ							SQSPKRYKLRDDFEKKMADFHKEEMDDQDK
DKAKGRKESEFDDEPKFMSKVIGANKNQE							DKAKGRKESEFDDEPKFMSKVIGANKNQEEE
							KSGKWEGLVYAPPGKEKQRKTEELEEESFPE
]		}				RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK
							AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS SFSITREAQVNVRMDSFDEDLARPSGLLAQER
							KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP
					·		SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT
	1		'				KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK
HGLAHDEMKSPREPGYKAEGKYKDDPVD] :				HGLAHDEMKSPREPGYKAEGKYKDDPVDLR
LDIERRKKHKERDLKRGKSRESVDSRDSSH				•			LDIERRKKHKERDLKRGKSRESVDSRDSSHSR
							ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS
SSSQSSHSYKAEEYTEETEEREESTTGFDKS	L	<u> </u>			L		SSSQSSHSYKAEEYTEETEEREESTTGFDKSRL

RGNYSGNNNNSNDDFQKRNREEEWDPFD	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTKDFVGPSERGGGRARGTFQFRARGRGWG RGNYSGNNNNNSNNDFQKRNREEEWDPEYT PKSKKYYLHDDREGEGSDKWVSRGRGRGAF	di di	nucleoti location correspo ng to fir amino a residue	in USSN 09/496			
Seq- uence 09/496 09/496 orrespondi ng to first amino acid residue of peptide sequence Properties sequence Open the sequence Open	Alast amino Cid residue Feptide Fuence M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTKDFVGPSERGGGRARGTFQFRARGRGWG RGNYSGNNNNNSNNDFQKRNREEEWDPEYT PKSKKYYLHDDREGEGSDKWVSRGRGRGAF	di d	ng to fir amino a residue	09/496	ı	peptide	nucl-
uence 914 amino acid residue of peptide sequence 7	cid residue f peptide f peptide quenc	i	ng to fir amino a residue	1	Į.	seq-	eotide
amino acid residue of peptide sequence p	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTKDFVGPSERGGGRARGTFQFRARGRGWG RGNYSGNNNNNSNNDFQKRNREEEWDPEYT PKSKKYYLHDDREGEGSDKWVSRGRGRGAF	i	amino a	914		uence	
residue of peptide sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTKDFVGPSERGGGRARGTFQFRARGRGWG RGNYSGNNNNNSNNDFQKRNREEEWDPEYT PKSKKYYLHDDREGEGSDKWVSRGRGRGAF		residue		ĺ		uence
peptide sequence	/=possible nucleotide deletion, \=possible nucleotide insertion GTKDFVGPSERGGGRARGTFQFRARGRGWG RGNYSGNNNNNSNNDFQKRNREEEWDPEYT PKSKKYYLHDDREGEGSDKWVSRGRGRGAF						i
	nucleotide insertion GTKDFVGPSERGGGRARGTFQFRARGRGWG RGNYSGNNNNNSNNDFQKRNREEEWDPEYT PKSKKYYLHDDREGEGSDKWVSRGRGRGAF	+			l		
GTKDFVGPSERGGRARGTFQFRARGRGW RGNYSGNNNNSNNDFQKRNREEEWDPES PKSKKYYLHDDREGEGSDKWVSRGRGRGRGPPKGRGFMFRKSSTSPKWAHDKFSGEGEGIDESGTENREEKDNIQPTTE 997 2347 A 8398 202 552 CPALGGRQDLQGTRLLWAHDSGVGGQKAI KQENLESLEATGREEEGGQGPPVTTKGVLL LLMAGLALQPGTALLCYSCKAQVSNEDCL VENCTQLGEQCWTARREWGDDSRQA 998 2348 A 8400 697 301 NPFSACTPGSCDSCSGRGRDLAFDSVWSTN MSDPRRPNKVLRYKPPPSECNPALDDPTPD MNLGMIFSMCGLMLKLKWCAWVAVYCSSFANSRSSEDTKQMMSSFMLSISAVVMSYLNPQPMTPPP 999 2349 A 8401 93 1126 ASASHITSGHLRCFPGSEGVGTMARCFSLVILTSIWTTRLLVQGSLRAEELSIQVSCRIMGIV VSKKANQQLNFTEAKEACRLLGLSLAGKDVETALKASFETCSYGWVGDGFVVISRISPNCGKNOVGVLIWKVPVSRQFAAYCYNSDDTTNSCPPEIITKDPIPTNTQTATQTTEFIVSDSTVASPYSTIPAPTTTPPAPASTSIPRRKKLICVVFMETSTIMSTETEPFVENKAAFKNEAGFCVPTALLVLALLFFGAAAGLGFCYVKRYVK.PFTNKNQKEMIETKVVKEEKANDSNPNEIKKTDKNPEESSSSEXTTMRCLEAEV 1000 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIF AAIFTADGNMISASTLMDILLMNDFKLVINN AYDVQCPKREKPSNEHTAMEHMKSLVHFFTILHLEESQKKREHHLLEKIDHLKEQLQPLQVKAGIEAHSEAKTSGLLWAGLALLSIQGC LAWLTVWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLC QMQVEELIERN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAFT GALLLRGSLLASGRAPRASSGLPRNTVV VPQQEAWVYERMGRFHRILEPGLNILIPVLIV	GTKDFVGPSERGGGRARGTFQFRARGRGWG RGNYSGNNNNNSNNDFQKRNREEEWDPEYT PKSKKYYLHDDREGEGSDKWVSRGRGRGAF	+			ĺ		
RGNYSGNNNNSNDIPGKRNREEEWDPPE PKSKKYY1HDDREGEGSDKWVSRGRGRGR PRGRGRFMFRKSSTSPKWAHDKFSGEGEI DDESGTENREEKDNIQPTTE DDESGTENREEKDNIQPTTE DDESGTENREEKDNIQPTTE CPALGGRQDLQGTRLLWAHDSGVGGCKAI KQENLESLEATGREEEGGGPPVTTKGVLL LLMAGLALQPGTALLCYSCKAQVSNEDCL VENCTQLGEQCWTARIREWGDDSRQA PSPSACTTPSCDSCSGGRGRDLAFDSVWSTN MSDPRRPNKVLRYKPPPSECNPALDDPTPD MNLIGMIFSMCGLMLKKWCAWVAVVCS SFANSRSSEDTKQMMSSFMLSISAVVMSVL NPQPMTPPW SASSHITSGHLRCFPGSEGVGTMARCFSLVI LTSIWTTRLLVQGSLRAEELSIQVSCRIMGII VSKKANQQLNFTEAKEACRLLGLSLAGKD VVETALKASFETCSYGWVGDGFVVISRISPN CGKNOVGVLIWKVPVSRQFAAYCYNSSDT TNSCPEIITTKDPINTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRKKLICV VFFMESTMISTETEPFVENKAAFKNEAAGF VPTALLVLALLFFGAAAGLGFCYVKRYVK PFTNKNQKEMIETKVVKEEKANDSNPNE KKIDKNPEESKSPSKITIMCLEAEV LOOO 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIF AAIFTADQNIMISASTLMDILLMNDFKLVINI AYPVQCPKRKPSNEHTAEMEHMKSLVHF FTILHLESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGC LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSK HPDVQQYNKLKEDLAKAKESLKQARHSLC QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRSLRRWEMLARAAF GALLLRGSLLASGRAPRASSGLPRNTVV VPQGEAWVVERMGRFHRILEPGLNILIPVLI VPQGEAWVVERMGRFHRILEPGLNILIPVLI VPQEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPQCEAWVVERMGRFHRILEPGLNILIPVLI VPCCEAWVERMGRFHRILEPGLNILIPVLI VPCCEAWVVERMGRFHRILEPGLNI	RGNYSGNNNNNSNNDFQKRNREEEWDPEYT PKSKKYYLHDDREGEGSDKWVSRGRGRGAF	- 1	sequenc	 	ļ		ļ
PKSKKYYLHDDREGEGSDKWVSRGRGRG/PRGRGFMFRKSSTSPKWAHDKFSGEGGE PRGRGFMFRKSSTSPKWAHDKFSGEGGE PRGRGFMFRKSSTSPKWAHDKFSGEGGE PRGRGFMFRKSSTSPKWAHDKFSGEGGE PRGRGFMFRKSSTSPKWAHDKFSGEGGE DDESGTENREEKDNIQPTTE LIMGLALOFICE CPALGGRQDLQGTRLLWAHDSGVGGQKAI KQENLESLEATGREEEGGGGPPVTTKGVLL LLMAGLALOFALLCYSCKAQVSNEDCL VENCTQLGEQCWTARIREWGDDSRQA P98	PKSKKYYLHDDREGEGSDKWVSRGRGRGAF	- 1	1	1	ļ		
PRGRGRFMFRKSSTSPKWAHDKFSGEGEI 1997 2347 A 8398 202 552 CPALGGRQDLQGTRLLWAHDSGVGGQKAI CPALGGRQDLQGTRLLWAHDSGVGGQKAI KQENLESLEATGREEGGQGPPVTTKGVLL LLMAGLALQPGTALLCYSCKAQVSNEDLO- VENCTQLGEQCWTTARREWGDDSRQA 998 2348 A 8400 697 301 NPFSACTFGSCDSCSGRGDLAFDSVWSTN MSDPRRPNKVLRYKPPPSECNPALDDPTPD MMLLGMIFSMCGLMLKLKWCAWVAVYCS SFANSRSSEDTKQMMSSFMLSISAVVMSYL NPQPMTPPW 999 2349 A 8401 93 1126 ASASHITSGHLRCFPGSEGVGTMARCFSLVI LTSIWTTRLLVQGSLRAEELSIQVSCRIMGIT VSKKANQQLNFTEAKEACRLLGLSLAGKD VSKKANQQLNFTEAKEACRLLGLSLAGKD VSKANQQLNFTEAKEACRLLGLSLAGKD VSKANQQLNFTEAKEACRLLGLSLAGKD VSKANQQLNFTEAKEACRLLGLSLAGKD VFTALKASFETTCSYGWVGDGFVVISRISPNI CGKNGVGVLIWKVPVSRQFAAYCYNSSDT TNSCIPELITIKDPIENTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRKKLICV VFMETSTMSTETEPFVENKAAFKNEAAGF VPTALLVLALLFFGAAAGLGFCYVKRYVK PFTRKENQGKMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV RERCQFVVKPMLSTVGSFLQDLQNEDKGIF AAIFTADGNMISASTLMDILLMNDFKLVIN AYDVQCFKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGG LAWLTWWVYSWDIMEPVTYFTIFANSMVI AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HPVQQYNKLKEDLAKAKESLKQARHSLC QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERFLRSSRWFRSILRRWEMLARAAF TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGREHRILEFGLNILIPVIJ		- 1					
DDESGTENREEKDNIQPTTE	I PRORUGREMENT STRUK WAHLIK ESCHEGELE				Ì		
997 2347 A 8398 202 552 CPALGGRQDLQGTRLLWAHDSGVGGQKAI KQENLESLEATGREEGGQGPPVTTKGVLL LLMAGLALOPGTALLCYSCKAQVSNEDCL VENCTQLGEQCWTARIREWGDDSRQA 998 2348 A 8400 697 301 NPPSACTPGSCDSCSGRGRDLAFDSVWSTN MSDPRRPNKVLRYKPPPSECNPALDDPTPD MNLGMIFSMCGLMLKLKWCAWVAVYCS FANSRSSEDTKQMMSSFMLSISAVVMSYL NPQPMTPPW 999 2349 A 8401 93 1126 ASSHITSGHLRCFPGSEGVGTMARCFSLVI LTSIWTTRLLVQGSLRAEELSIQVSCRIMGIT VSKKANQQLNFTEAKEACRLLGLSLAGKD VETALKASFETCSYGWVGDGFVVISRISPNI CCKNGVGVLIWKVPVSRQFAAYCYNSSDT TNSCIPEIITTKDPIFNTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTMSTETEPFVENKAAFKNEAAGFQ VPTALLVLALLFFGAAAGLGFCYVKRYVK PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 777		- 1					
SQENLESLEATGREEEGGOGPPVTTKCVLL LLMAGLALOPGTALLCYSCKAQVSNEDCL VENCTQLGEQCWTARIREWGDDSRQA 998 2348 A 8400 697 301 NPPSACTPGSCDSCSGRGRDLAFDSVWSTN MSDPRRPNKVLRYKPPPSECNPALDDPTPD MNLLGMIFSMCGLMLKLKWCAWVAVCS SFANSRSEDTKQMMSSFMLSISAVVMSYL NPQPMTPPW SPANSRSEDTKQMMSSFMLSISAVVMSYL NPQPMTPPW ASASHTSGFILRCPPGSEGVGTMARCFSLVI LTSIWTTRILVQGSLRAEELSIQVSCRIMGI VSKKANQQLNFTEAKEACRLLGLSLAGKD VETALKASFETCSYGWVGDGFVVISRISPNI CCKNGVGVLIWKVPVSRQFAAYCYNSSDT TNSCIPEIITTKDPIFNTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTMSTETEFFVENKAAFKNEAAGFC VPTALLVLALLFFGAAAGLGFCYVKRYVK PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV VFMETSTWSTFGFLQDLQNEDKGIR AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKESNEHTIAEMEHMKSLVHF FTILHLEESQKKRFHHILEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGC LAWLTWWYYSWDIMEPVTYFITANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKSELKQARHSLC QMQVEELNEKN QMQVEELNEKN VGFWERPJLRSSRWFRRSLRRWEMLARAAFT TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI		十	202	8398	A	2347	997
998 2348 A 8400 697 301 NPPSACTPGSCQSCSGRGRDLAFDSVWSTN MSDPRRPNKVLRYKPPPSECNPALDDPTPD MNLLGMIFSMCGLMLKLKWCAWVAVYCS SFANSRSSEDTKQMMSSFMLSISAVVMSYL NPOPMTPPW 999 2349 A 8401 93 1126 ASASHITSGHLRCFPGSEGVGTMARCFSLVI LTSIWTTRLLVQGSLRAEELSIQVSCRIMGII VSKKANQQLNFTEAKEACRLLGLSLAGKD VETALKASFETCSYGWVGDGFVVISRISPNI CGKNGVGVLIWKVPVSRQFAAYCYNSSDT TNSCIPEIITTKDPIFNTQTATQTTEFFVSDST VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTMSTETEPFVENKAAFKNEAAGFC VPTALLVLALLFFGAAAGLGFCYVKRYVK PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIF AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGC LAWLTWWVYSWDIMEPVTYFITFANSMVI AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLC QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSSRERRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFFHILLEPGLNILIPVLI		-		10070			***
VENCTQLGEQCWTARIREWGDDSRQA							
MSDPRRPNKVLRYKPPPSECNPALDDPTPD MNILIGMIFSMCGILMI.KLKWCAWVAVYCS SFANSRSSEDTKQMMSSFMLSISAVVMSYL NPQPMTPPW 999 2349 A 8401 93 1126 ASASHITSGHLRCFPGSEGVGTMARCFSLVI LTSIWTTRLLVQGSLRAEELSIQVSCRIMGI VSKKANQQLNFTEAKEACRLLGLSLAGKD VETALKASFETCSYGWVGDGFVVISRISPNI CGKNGVGVLIWKVPVSRQFAAYCYNSSDT TNSCIPEIITTKDPIFNTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTMSTETEPFVENKAAFKNEAAGFG VPTALLVLALLFFGAAAGLGFCYVKRYVK PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKKRFHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGGIALLSIQGG LAWLTWVYSWDIMEPVTYFITFANSMVI AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLC QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAF TGALLLRGSLLASGRAPRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI		- 1					
MNILGMIFSMCGLMLKLKWCAWVAVYCS SFANSRSSEDTKQMMSSFMLSISAVVMSYL NPQPMTPPW 999 2349 A 8401 93 1126 ASASHITSGHLRCFPGSEGVGTMARCFSLVI LTSIWTTRLLVQGSLRAEELSIQVSCRIMGIT VSKKANQQLNFTEAKEACRLLGLSLAGKD VETALKASFETCSYGWVGDGFVVISRISPNI CGKNGVGVLIWKVPVSRQFAAYCYNSSDST TNSCIPEIITTKDPIFNTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTMSTETEPFVENKAAFKNEAAGFC VPTALLVLALLFFGAAAGLGFCYVKRYVK PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGG LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLC QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAF TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI	NPPSACTPGSCDSCSGRGRDLAFDSVWSTNN	_	697	8400	A	2348	998
SFANSRSSEDTKQMMSSFMLSISAVVMSYL NPQPMTPPW 999 2349 A 8401 93 1126 ASASHITSGHLRCFPGSEGVGTMARCFSLVI LTSIWTTRLLVQGSLRAEELSIQVSCRIMGII VSKKANQQLINFTEAKEACRLLGLSLAGKD VETALKASFETCSYGWVGDGFVVISRISPNI COKNGVGVLIWKVPVSRQFAAYCYNSSDT TNSCIPEIITTKDPIFNTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTMSTETEPFVENKAAFKNEAAGFC VPTALLVLALLFFGAAAGLGFCYVKRYVK. PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGC LAWLTWWVYSWDIMEPVTYFITFANSMVI AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLC QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAF TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI	MSDPRRPNKVLRYKPPPSECNPALDDPTPDY	- 1			1		
NPQPMTPPW	MNLLGMIFSMCGLMLKLKWCAWVAVYCSFI				İ	,	
999 2349 A 8401 93 1126 ASASHITSGHLRCFPGSEGVGTMARCFSLVI LTSIWTTRLLVQGSLRAEELSIQVSCRIMGIT VSKKANQQLNFTEAKEACRLLGLSLAGKD VETALKASFETCSYGWVGDGFVVISRISPNI CGKNGVGVLIWKVPVSRQFAAYCYNSSDT TNSCIPEIITTKDPIFNTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTIMSTETEPFVENKAAFKNEAAGFC VPFALLVLALLFFGAAAGLGFCYVKRYVK PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 7777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIF AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGG LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLG QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI	SFANSRSSEDTKQMMSSFMLSISAVVMSYLQ	- 1		1	ł		
LTSIWTTRLLVQGSLRAEELSIQVSCRIMGIT VSKKANQQLNFTEAKEACRLLGLSLAGKD VETALKASFETCSYGWVGDGFVVISRISPNI CGKNGVGVLIWKVPVSRQFAAYCYNSSDT TNSCIPEIITTKDPIFNTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTMSTETEPFVENKAAFKNEAAGFQ VPTALLVLALLFFGAAAGLGFCYVKRYVK. PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGG LAWLTWWYYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLG QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI							
VSKKANQQLNFTEAKEACRLLGLSLAGKD VETALKASFETCSYGWVGDGFVVISRISPNI CGKNGVGVLIWKVPVSRQFAAYCYNSSDT TNSCIPEIITTKDPIFNTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTMSTETEPFVENKAAFKNEAAGFC VPTALLVLALLFFGAAAGLGFCYVKRYVK. PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGO LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLO QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI		T	93	8401	A	2349	999
VETALKASFETCSYGWVGDGFVVISRISPNI CGKNGVGVLIWKVPVSRQFAAYCYNSSDT TNSCIPEIITTKDPIFNTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTMSTETEPFVENKAAFKNEAAGFG VPTALLVLALLFFGAAAGLGFCYVKRYVK PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGG LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLC QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAF TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI	LTSIWTTRLLVQGSLRAEELSIQVSCRIMGITL		1		j]]
CGKNGVGVLIWKVPVSRQFAAYCYNSSDT TNSCIPEIITTKDPIFNTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTMSTETEPFVENKAAFKNEAAGFO VPTALLVLALLFFGAAAGLGFCYVKRYVK PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHILEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGO LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLO QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI							
TNSCIPEIITTKDPIFNTQTATQTTEFIVSDST VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTMSTETEPFVENKAAFKNEAAGFO VPTALLVLALLFFGAAAGLGFCYVKRYVK PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGO LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLO QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI							İ
VASPYSTIPAPTTTPPAPASTSIPRRKKLICV VFMETSTMSTETEPFVENKAAFKNEAAGFC VPTALLVLALLFFGAAAGLGFCYVKRYVK PFTNKNQQKEMETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGC LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLC QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI		- 1			İ		1
VFMETSTMSTETEPFVENKAAFKNEAAGFG VPTALLVLALLFFGAAAGLGFCYVKRYVK. PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHILESQKKREHHILLEKIDHILKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGG LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLG QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI							
VPTALLVLALLFFGAAAGLGFCYVKRYVK. PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGO LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLO QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI			1	i			ŀ
PFTNKNQQKEMIETKVVKEEKANDSNPNEI KKTDKNPEESKSPSKTTMRCLEAEV 1000 2350 A 8406 2 7777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGO LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLO QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI					}	}	
KKTDKNPESKSPSKTTMRCLEAEV					· .		
1000 2350 A 8406 2 777 KERCQFVVKPMLSTVGSFLQDLQNEDKGIK AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGO LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLO QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAF TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI	1	- 1			ŀ	Ì	
AAIFTADGNMISASTLMDILLMNDFKLVINI AYDVQCPKREKPSNEHTAEMEHMKSLVHF FTILHILEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGO LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLO QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI			+-	8406	A	2350	1000
AYDVQCPKREKPSNEHTAEMEHMKSLVHIF FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGG LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLG QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI		ĺ	1-	0.00	1	1 2000	1000
FTILHLEESQKKREHHLLEKIDHLKEQLQPL QVKAGIEAHSEAKTSGLLWAGLALLSIQGG LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLG QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI							
QVKAGIEAHSEAKTSGLLWAGLALLSIQGO LAWLTWWVYSWDIMEPVTYFITFANSMVF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLO QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI			1			1	'
AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQ HFDVQQYNKLKEDLAKAKESLKQARHSLO QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI	QVKAGIEAHSEAKTSGLLWAGLALLSIQGGA		ļ				
HFDVQQYNKLKEDLAKAKESLKQARHSLC QMQVEELNEKN 1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI	LAWLTWWVYSWDIMEPVTYFITFANSMVFF	- [
1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI	AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQQ	- }			İ		
1001 2351 A 8410 1400 264 VGFWERPLRSSRWFRRSLRRWEMLARAAR TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI	HFDVQQYNKLKEDLAKAKESLKQARHSLCL	- 1				ļ	
TGALLLRGSLLASGRAPRRASSGLPRNTVV VPQQEAWVVERMGRFHRILEPGLNILIPVLI		لــــــــــــــــــــــــــــــــــــــ			<u> </u>		
VPQQEAWVVERMGRFHRILEPGLNILIPVLI			1400	8410	_A	2351	1001
	TGALLLRGSLLASGRAPRRASSGLPRNTVVLF				ļ		1
		1	J		1]	
	IRYVQSLKEIVINVPEQSAVTLDNVTLQIDGV	-		1			
	LYLRIMDPYKASYGVEDPEYAVTQLAQTTM			1		1	
	RSELGKLSLDKVFRERESLNASIVDAINQAAD	1					
	CWGIRCLRYEIKDIHVPPRVKESMQMQVEAE RRKRATVLESEGTRESAINVAEGKKOAOILAS			}	ł	}	1
	. EAEKAEQINQAAGEASAVLAKAKAKAEAIRI		1		1		[
	LAAALTQHNGDAAASLTVAEQYVSAFSKLA		ı			1	
	KDSNTILLPSNPGDVTSMVAQAMGVYGALT		1				
	I WOULTING OUT ON A LOIM AUGUMO A LOURI			1		1	[
	KAPVPGTPDSLSSGSSRÐVQGTDASLDEELDR		134	8421	A	2352	1002
	KAPVPGTPDSLSSGSSRÐVQGTDASLDEELDR VKMS	-+	1		1		
	KAPVPGTPDSLSSGSSRÐVQGTDASLDEELDR VKMS 41 NRENLLESRMMDPCSVGVQLRTTNECHKTY	十		i	1		
	KAPVPGTPDSLSSGSSRÐVQGTDASLDEELDR VKMS 41 NRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG	1			ł	[l .
	KAPVPGTPDSLSSGSSRÐVQGTDASLDEELDR VKMS 41 NRENLLESRMMDPCSVGVQLRTTNECHKTY					1	1
ALRSLQFTNPGRQTEFAPETGKREKRRLTK	KAPVPGTPDSLSSGSSRÐVQGTDASLDEELDR VKMS NRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC				1		
ATAGSDRQVIPAKSKVYDSQGLLIFSGMDL	KAPVPGTPDSLSSGSSRDVQGTDASLDEELDR VKMS INRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC PKCTQIINGSVDVDTEDRQKRKPESDGRTAK ALRSLQFTNPGRQTEFAPETGKREKRRLTKN					1	i
	KAPVPGTPDSLSSGSSRDVQGTDASLDEELDR VKMS INRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC PKCTQIINGSVDVDTEDRQKRKPESDGRTAK ALRSLQFTNPGRQTEFAPETGKREKRRLTKN						
WLYEQIEIEGGEIIHNKHAG	KAPVPGTPDSLSSGSSRÐVQGTDASLDEELDR VKMS INRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC PKCTQIINGSVDVDTEDRQKRKPESDGRTAK						
	KAPVPGTPDSLSSGSSRDVQGTDASLDEELDR VKMS INRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC PKCTQIINGSVDVDTEDRQKRKPESDGRTAK ALRSLQFTNPGRQTEFAPETGKREKRRLTKN ATAGSDRQVIPAKSKVYDSQGLLIFSGMDLC DCLDEDCLGCFYACPACGSTKCGAECRCDRK						
	KAPVPGTPDSLSSGSSRDVQGTDASLDEELDR VKMS INRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC PKCTQIINGSVDVDTEDRQKRKPESDGRTAK ALRSLQFTNPGRQTEFAPETGKREKRRLTKN ATAGSDRQVIPAKSKVYDSQGLLIFSGMDLC DCLDEDCLGCFYACPACGSTKCGAECRCDRK WLYEQIEIEGGEIIHNKHAG TEWGLSGSCPGCSPLEPGSRGRGAAAWRILR		3	8427	A	2353	1003
MHPAVFLSLPDLRCSLLLLVTWVFTPVTTE	KAPVPGTPDSLSSGSSRÐVQGTDASLDEELDR VKMS INRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC PKCTQIINGSVDVDTEDRQKRKPESDGRTAK ALRSLQFTNPGRQTEFAPETGKREKRRLTKN ATAGSDRQVIPAKSKVYDSQGLLIFSGMDLC DCLDEDCLGCFYACPACGSTKCGAECRCDRK WLYEQIEIEGGEIIHNKHAG TEWGLSGSCPGCSPLEPGSRGRGAAAWRILR CRRLPEPSPFLTQPNLAQSQPPAPVPVTDPSVT		3	8427	A	2353	1003

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino soid conservation C. Contribution
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in NO:	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
цепсе		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
101100			^^-	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1	1	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
		 		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		SLDTENIDEILNNADVALVNFYADWCRFSQM
ĺ		Í	İ			LHPIFEEASDVIKEEFPNENQVVFARVDCDQH
						SDIAQRYRISKYPTLKLFRNGMMMKREYRGQ
		ļ				RSVKALADYIRQQKSDPIQEIRDLAEITTLDRS
			ŀ			KRNIIGYFEQKDSDNYRVFERVANILHDDCAF
Ì		ĺ	ľ			LSAFGDVSKPERYSGDNIIYKPPGHSAPDMVY
						LGAMTNFDVTYNWIQDKCVPLVREITFENGE
		j]			ELTEEGLPFLILFHMKEDTESLEIFQNEVARQL
1		1				ISEKGTINFLHADCDKFRHPLLHIQKTPADCP
		1	1			VIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDL
}		1	l			HSGKLHREFHHGPDPTDTAPGEQAQDVASSP
		L				PESSFQKLAPSEYRYTLLRDRDEL
1004	2354.	A	8432	910	387	GLSRKLRAGFLPGFCRVSPCGSWVVETLVKM
1]	1			ACAAARSPADQDRFICIYPAYLNNKKTIAEGR
			1			RIPISKAVENPTATEIQDVCSAVGLNVFLEKN
		1				KMYSREWNRDVQYRGRVRVQLKQEDGSLC
1		[[LVQFPSRKSVMLYAAEMIPKLKTRTQKTGGA
1005	2255	- <u>-</u>	0452	-	520	DQSLQQGEGSKKGKGKKKK
1002	2355	A	8453	90	530	QSHETKMQSGTHWRVLGLCLLSVGVWGQD
						GNEEMGGITQTPYKVSISGTTVILTCPQYPGSE
		1	1			ILWQHNDKNIGGDEDDKNIGSDEDHLSLKEF
	Ì	1		:		SELEQSGYYVCYPRGSKPEDANFYLYLRARG
1006	2356	A	8458	3	307	NPGLQNRYHRLFREDHSKGHSQ
1000	2330	^	07.70	,	307	AVQRIRHEMNIFRLTGDLSHLAAIVILLLKIW KTRSCAGISGKSQLLFALVFTTRYLDLFTSFIS
		[LYNTSMKVWYAIHRNVFHLQCTGLWTLNLC
		1	1			QLCIFN TAIHRIVVIALQCIGEWILLIC
1007	2357	A	8459	43	553	GAGAGGDWAAMDKLKKVLSGQDTEDRSGL .
1]					SEVVEASSLSWSTRIKGFIACFAIGILCSLLGT
1		1			1	VLLWVPRKGLHLFAVFYTFGNIASIGSTIFLM
						GPVKQLKRMFEPTRLIATIMVLLCFALTLCSA
1]	1			FWWHNKGLALIFCILQSLALTWYSLSFIPFAR
		<u></u>				DAVKKCFAVCLA
1008	2358	A	8462	487	150	AQDIRSVHSLGQKSTFVKHFRTLSHLHGLPDP
1		1				PPHWPPQERSPPSHPCMPSHRPQIPQLSNSGPS
		İ			!	DPRWGCVGPSMPTSTCLPGAVEASTTKASLP
		<u> </u>				KCPVDSSLPTPEACFL
1009	2359	Α	8465	134	954	ETRVKTSLELLRTQLEPTGTVGNTIMTSQPVP
		}				NETIIVLPSNVINFSQAEKPEPTNQGQDSLKKH
		1			,	LHAEIKVIGTIQILCGMMVLSLGIILASASFSPN
						FTQVTSTLLNSAYPFIGPFFFIISGSLSIATEKRL
İ		1			1	TKLLVHSSLVGSILSALSALVGFIILSVKQATL
		1	[NPASLQCELDKNNIPTRSYVSYFYHDSLYTTD
		1				CYTAKASLAGTLSLMLICTLLEFCLAVLTAVL
		ĺ	[RWKQAYSDFPGSVLFLPHSYIGNSGMSSKMT
1010	2360	 	0460	-	472	HDCGYEELLTS
1010	2300	A	8468	2	473	KYRYRPYPVMRKICQVGPAGLAFILNISPVA
		1				HRVALCHLAGCQEQAAWYHTLQILFFLVSAY
	<u>'</u>					FFSCPVPEKYFPGSCDIVGHGHQIFHAFLSICT
}		l				LSQLEAILLDYQGRQEIFLQRHGPLSVHMACL SFFFLAACSAATAALLRHKVKARLTKKDS
1011	2361	A	8478	5	409	TELSQLEKAHPPADMGRRKSKRKPPPKKKMT
1011	2301	1^	04/0		407	GTLETQFTCPFCNHEKSCDVKMDRARNTGVI
1			}			SCTVCLEEFQTPITCILGNLGFFQRVGRGLESG
] .			PCSSGPLCALVQGQSRPEEQVPPSDFCGVRRC
]				RAGFOCO
1012	2362	A	8481	2810	1652	RTSTQKWQSVFNDSQEHLERFYCNPENDRM
		l '`	0.01	2010	1072	RMKYGGQEFWADLNAMNVYETTEFDQLRR
						LSTPPSSNVNSIYHTVWKFFCRDHFGWREYPE
<u> </u>	<u> </u>					DOILL DOILL THAT CADIFO WKEIFE

- OPO	L CO CO	1 1 4 - 2	CEO	Dendier 1	Dun alternation	L Amino anid commune (A - Alexin - C-C
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl- eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	GOILE	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l Boilea	1	1	1 7 4 7	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	İ	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
		1		sequence		nucleotide insertion
 		 				SVIRLIEEANSRGLKEVRFMMWNNHYILHNS
Ī	ļ	1			1	FFRREIKRRPLFRSCFILLPYLQTLGGVPTQAP
1		}			}	PPLEATSSSQIICPDGVTSANFYPETWVYMHP
ĺ	[1				SQDFIQVPVSAEDKSYRIIYNLFHKTVPEFKYR
ł					į	ILQILRVQNQFLWEKYKRKKEYMNRKMFGR
	1	ŀ	1			DRINERHLFHGTSQDVVDGICKHNFDPRVCG
ļ	1	ł	1	1]	KHATMFGQGSYFAKKASYSHNFSKKSSKGV
,		ļ	1			HFMFLAKVLTGRYTMGSHGMRRPPPVNPGS
		l]	VTSDLYDSCVDNFFEPQIFVIFNDDQSYPYFVI
			L			QYEEVSNTVSI
1013	2363	A	8488	2	517	IENCRTRLRQAWHEVCGNKMAAPIPQGFSCL
						SRFLGWWFRQPVLVTQSAAIVPVRTKKRFTP
]	1	1				PIYQPKFKTEKEFMQHARKAGLVIPPEKSDRS
		l		ļ		IHLACTAGIFDAYVPPEGDARISSLSKEGLIER
1	1	l				TERMKKTMASQVSIRRIKDYDANFKIKDFPE
						KAKDIFIEGSPLY
1014	2364	A	8501	363	17	YIRTGYVYICIIYAQLMYTYYIRTAYVYICILY
	1					AQLMYTYVLYTHSLCIHMYSIRTAYVYICIIY
						AQIMYTYVFYTHRLCIHMYSIRTDYVYICILY
		<u> </u>				AQLMYTYVFYTHSYMSDE
1015	2365	A	8504	3	2190	NSSEHFSQAPQRLSFYSWYGSARLFRFRVPPD
	1	ŀ				AVLLRWLLQVSRESGAACTDAEITVHFRSGA
	1			ŀ		PPVINPLGTSFPDDTAVQPSFQVGVPLSTTPRS
[1	f .			NASVNVSHPAPGDWFVAAHLPPSSQKIELKG
]	!	1				LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ
l		1				TLLSHPSYLKVFVPDYTRELLLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC
		ļ				RLLLPSPPWDRWLQVTAESLVGPLGTVAFSA
·	1	1				VAALTACRPRSVTIQPLLQSSQNQSFNASSGL
		1		l	ļ	LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED
	1	ł		ł	ł	MDVVSVHFOPLDRVSVRVCSDTPSVMRLRL
	1					NTGMDSGGSLTISLRANKTEMRNETVVVACV
	[j		ļ	NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR
	1	1				RANLIIPYPETDNWYLSLQLMCPENAEDCEQ
	1					AVVHVETTLYLVPCLNDCGPYGQCLLLRRHS
		1				YLYASCSCKAGWRGWSCTDNSTAQTVAQQR
} .		1	}	1	}	AATLLLTLSNLMFLAPIAVSVRRFFLVEASVÝ
'	1	1			1	AYTMFFSTFYHACDQPGEAVLCILSYDTLQY
	1	1				CDFLGSGAAIWVTILCMARLKTVLKYVLFLL
ſ	1	ĺ		{	ĺ	GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM
	1		1	!		ASMWAYRCGHRRQCYPTSWQRWAFYLLPG
				1		VSMASVGIAIYTSMMTSDNYYYTHSIWHILL
						AGSAALLLPPPDQPAEPWACSQKFPCHYQIC
L			<u></u>		L	KNDREELYAVT
1016	2366	A	8511	1	453	KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK
		1	1	Į		KGGEKKKGRSAINEVVTREYTINIHKRIHGVG
		1				FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL
			1	1		NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP
						NKLYTLVTYVPVTTFKNLQTVNVDEN
1017	2367	A	8513	54	1196	LERTPASADMAWTKYQLFLAGLMLVTGSINT
	[}				LSAKWADNFMAEGCGGSKEHSFQHPFLQAV
	[1	GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ
1	1		1	ĺ		QPFNPLLFLPPALCDMTGTSLMYVALNMTSA
						SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL
		1	1		-	GILATIAGLVVVGLADLLSKHDSQHKLSEVIT
	[Ī	ŀ		GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR
	l	l	1	Í	l	AVGTEGLFGFVILSLLLVPMYYIPAGSFSGNP
	1					RGTLEDALDAFCQVGQQPLIAVALLGNISSIA
i	I	1	1	<u>L</u>	i	FFNFAGISVTKELSATTRMVLDSLRTVVIWAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SLALGWEAFHALQILGFLILLIGTALYNGLHR PLLGRLSRGRPLAEESEQERLLGGTRTPINDA S
1018	2368	A	8518	324	694	SPFWTEKRRMEKPLFPLVPLHWFGFGYTALV VSGGIVGYVKTGSVPSLAAGLLFGSLAGLGA YQLYQDPRNVWGFLAATSVTFVGVMGMRS YYYGKFMPVGLIAGASLLMAAKVGVRMLM TSD
1019	2369	A	8526	2	1787	VSAAAVNMEPPDAPAQARGAPRLLLLAVLL AAHPDAQAEVRLSVPPLVEVMRGKSVILDCT PTGTHDHYMLEWFLTDRSGARPRLASAEMQ GSELQVTMHDTRGRSPPYQLDSQGRLVLAEA QVGDERDYVCVVRAGAAGTAEAAARLNVF AKPEATEVSPNKGTLSVMEDSAQEIATSNSRN GNPAPKITWYRNGQRLEVPVEMNPEGYMTS RTVREASGLLSLTSTLYLRLRKDDRDASFHC AAHYSLPEGRHGRLDSPTFHLTLHYPTEHVQ FWVGSPSTPAGWVREGDTVQLLCRGDGSPSP EYTLFRLQDEQEEVLNVNLEGNLTLEGVTRG QSGTYGCRVEDYDAADDVQLSKTLELRVAY LDPLELSEGKVLSLPLNSRAVVNCSVHGLPTP ALRWTKDSTPLGDGPMLSLSSITFDSNGTYVC EASLPTVPVLSRTQNFTLLVQGSPELKTAEIEP KADGSWREGDEVTLICSARGHPDPKLSWSQL GGSPAEPIPGRQGWVSSSLTLKVTSALSRDGI SCEASNPHGNKRHVFHFGTVSPQTSQAGVAV MAVAVSVGLLLLVVAVFYCVRRKGGPCCRQ RREKGAP
1020	2370	A	8530	2	1200	PRVRLLRPSRSRSCRGLLSTRAPGPSPFRSLHS SPLLPHAMKSPFYRCQNITTSVEKGNSAVMGG VLFSTGLLGNLLALGLLARSGLGWCSRRPLR PLPSVFYMLVCGLTVTDLLGKCLLSPVVLAA YAQNRSLRVLAPALDNSLCQAFAFFMSFFGL SSTLQLLAMALECWLSLGHPFFYRRHITLRLG ALVAPVVSAFSLAFCALPFMGFGKFVQYCPG TWCFIQMVHEEGSLSVLGYSVLYSSLMALLV LATVLCNLGAMRNLYAMHRRLQRHPRSCTR DCAEPRADGREASPQPLEELDHLLLLALMTV LFTMCSLPVIYRAYYGAFKDVKEKNRTSEEA EDLRALRFLSVISIVDPWIFIIFRSPVFRIFFHKI FIRPLRYRSRCSNSTNMESSL
1021	2371	A	8536	1	237	RRGEIDMATEGDVELELETETSGPERPPEKPR KHDSGAADLERVTDYAEEKEIQSSNLETAMS VIGDRRSREQKAKQER
1022	2372	À	8537	94	541	RKERRRRRMEAVVFVFSLLDCCALIFLSV YFIITLSDLECDYINARSCCSKLNKWVIPELIG HTIVTVLLLMSLHWFIFLLNLPVATWNIYRYI MVPSGNMGVFDPTEIHNRGQLKSHMKEAMI KLGFHLLCFFMYLYSMILALIND
1023	2373	A	8540	26	431	RMMKCPQALLAIFWLLLSWVSSEDKVVQSPL SLVVHEGDTVTLNCSYEVTNFRSLLWYKQEK KAPTFLFMLTSSGIEKKSGRLSSILDKKELSSIL NITATQTGDSAIYLCAVEAQCSLVTCSLYSNS TAEALQL
1024	2374	A	8544	1731	743	GVRLRYSPIAVVMVGEAGRDLRRRRAVAVT AEKMAVLAPLIALVYSVPRLSRWLAQPYYLL SALLSAAFLLVRKLPPLCHGLPTQREDGNPCD FDWREVEILMFLSAIVMMKNRRSITVEQHIGN IFMFSKVANTILFFRLDIRMGLLYITLCIVFLM

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
4000	ļ			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	ŀ		1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	İ	1		peptide	sequence	/=possible nucleotide deletion, \=possible
	ŀ	i	1	1		
		<u> </u>	<u> </u>	sequence		nucleotide insertion
l	Į	Ì	ł	ł		TCKPPLYMGPEYIKYFNDKTIDEELERDKRVT
	ł					WIVEFFANWSNDCQSFAPIYADLSLKYNCTG
	Į.	ł	Į.	}		LNFGKVDVGRYTDVSTRYKVSTSPLTKQLPT
	1		1]		LILFQGGKEAMRRPQIDKKGRAVSWTFSEEN
	1		l			VIREFNLNELYQRAKKLSKAGDNIPEEQPVAS
l	1	i	1	İ		TPTTVSDGENKKDK
1025	2375	Ā	8546	2194	1707	TVSFHKTMASLKCSTVVCVICLEKPKYRCPA
1025	23/3	^	6540	2194	1707	CRVPYCSVVCFRKHKEQCNPETRPVEKKIRS
		1	ł		Ì	
		İ	ľ			ALPTKTVKPVENKDDDDSIADFLNSDEEEDR
	1	ĺ	1	Í		VSLQNLKNLGESATLRSLLLNPHLRQLMVNL
1						DQGEDKAKLMRAYMQEPLFVEFADCCLGIV
]				EPSQNEES
1026	2376	A	8547	1078	594	VGMELPAVNLKVILLGHWLLTTWGCIVFSGS
]]]		YAWANFTILALGVWAVAQRDSIDAISMFLGG
						LLATIFLDIVHISIFYPRVSLTDTGRFGVGMAIL
	1]			ł	SLLLKPLSCCFVYHMYRERGGELLVHTGFLG
		1	1		1	SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR
ļ	1	1	Į.			
						GY
1027	2377	A	8557	1	340	DFLGPASPQEEGGSESSTMTELETAMGMIIDV
		1	1			FSRYSGSEGSTQTLTKGELKVLMEKELPGFLQ
			1	l		SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF
1			1	ł		VAAITSACHKYFEKAGLK
1028	2378	A	8569	20	963	KMAATLGPLGSWQQWRRCLSARDGSRRLLL
1020	1270	1	0505	20	303	LLLLGSGQGPQQVGAGQTFEYLKREHSLSKP
		ł				YQGEAPRPCFLRDWELQVHFKIHGQGKKNL
1	i	1				
i		1				HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG
j	1		1	l		VFVDTYPNEEKQQERVFPYISAMVNNGSLSY
1		1	1	1		DHERDGRPTELGGCTAIVRNLHYDTFLVIRY
		1	1			VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
1				1		YYFGTSSITGDLSDNHDVISLKLFELTVERTPE
ł						EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL
ŀ	1	1		}	l	ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK
		l	1	į.		RFY
1029	2379	A	8572	<u> </u>	578	AAAASHRSRARSRPRRVSSGPAPRRAQSSAG
1029	23/3	1^	8372	1 '	378	RVASGLDSAPLCTMARALCRLPRRGLWLLLA
ł	1			,		
1	1	ł	1	ł	1	HHLFMTTACQEANYGALLRELCLTQFQVDM
	1	1		1	1	EAVGETLWCDWGRTIRSYRELADCTWHMAE
	1	1	1	1		KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR
1				1	1	AVRDPPGSILYPFIVVPITVTLLVTALVVWQS
		1		1	1	KRTEGIV
1030	2380	A	8574	1352	372	DSSTVKGGSESRHLCLIPDLKGKARTREASSG
	1	1		1		SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG
1		1			1	THLTITQALRQPLHRAPLLPGQLCWSPRPLEK
	1	1	I	1		
			1		1	NKAMGRPLLLPLLLLQPPAFLQPGGSTGSGP
1		1	1	[[SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL
1				1	1	AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY
1		1			Į.	VNRLFLNWTEGQESGFLRISNLRKEDQSVYF
		!	1			CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT
		1			ļ	TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT
		1	1	1	!	AIRVALAVAVLKTVILGLLCLLLLWWRRRKG
		1			1	SRAPSSDF
1031	2381	A	8580	905	340	RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL
1021	اه ت	^	0200	303	340	
]	1]]]	AFDDFQESCAMMWQKYAGSRRSMPLGARIL
	}	1			1	FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV
	1	1			1	EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI
		1	1		1	VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW
1					1	HLDEVFLELKDGQQIPVFKLSGENGDEVKKE
1032	2382	1A	8593	2558	961	RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR
						WAAGALGVAGLLCAVLGAVMIVMVPSLIKQ
L				<u> </u>	L	"" TO TO THE TANGET OF THE TANK A DOME OF

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QVLKNVRIDPSSLSFNMWKEIPIFYLSVYFFD VMNPSEILKGEKPQVRERGPYVYREFRHKSNI TFNNNDTVSFLEYRTFQFQPSKSHGSESDYIV MPNILVLGAAVMMENKPMTLKLIMTLAFTTL
1033	2383	A	8595	595	767	GERAFMNRTVGEIMWGYKDPLVNLINKYFP GMFPFKDKFGLFAELNNSDSGLFTGFTGVQNI SRIHLVDKWNGLSKVDFWHSDQCNMINGTS GQMWPPFMTPESSLEFYSPEACRSMKLMYKE SGVFEGIPTYRFVAPKTLFANGSIYPPNEGFCP CLESGIQNVSTCRFSAPLFLSHPHFLNADPVL AEAVTGLHPNQEAHSLFLDIHPVTGIPMNCSV KLQLSLYMKSVAGIGQTGKIEPVVLPLLWFA ESGAMEGETLHTFYTQLVLMPKVMHYAQYV LLALGCVLLLVPVICQIRSQEKCYLFWSSSKK GSKDKEAIQAYSESLMTSAPKGSVLQEAKL AHLPDTLLLPPHSPTVPTPKSFQCSQKACFSRS
			<u> </u>			FCLLLSLVSSSLVSLSLCPPLTQA
1034	2384	A	8597	640	164	VTTSCIIPFAFGLGVRASERLAEIDMPYLLKYQ PMMQTIGQKYCMDPAVIAGVLSRKSPGDKIL VNMGDRTSMVQDPGSQAPTSWISESQVFQTT EVLTTRITELQRRFPTWTPDQYLRGGLCAYSG GAGYVRSSQDLSCDFCNDVLARAKYLKRHG F
1035	2385	A	8603	936	204	AMASTLEYSPSPLRRLVGPAAGFSRAARADL SWDPMAFFTGLWGPFTCVSRVLSHHCFSTTG SLSAIQKMTRVRVVDNSALGNSPYHRAPRCI HVYKKNGVGKVGDQILLAIKGQKKKALIVG HCMPGPRMTPRFDSNNVVLIEDNGNPVGTRI KTPIPTSLRKREGEYSKVLAIAQNFV
1036	2386	A	8606	1	562	PTRAHSFDLCCSPCRRRLLGREEAGEEPTSPV TQYLQPRSPEECKMFACAKLACTPSLIRAGSR VAYRPISASVLSRPEASRTGEGSTVFNGAQNG VSQLIQREFQTSAISRDIDTAAKFIGAGAATVG VAGSGAGIGTVFGSLIIGYARNPSLKQQLFSY AILGFALSEAMGLFCLMVAFLILFAM
1037	2387	A	8615	2	2364	SPGPSLPESAESLDGSQEDKPRGSCAEPTFTDT GMVAHINNSRLKAKGVGQHDNAQNFGNQSF EELRAACLRKGELFEDPLFPAEPSSLGFKDLG PNSKNVQNISWQRPKDIINNPLFIMDGISPTDI CQGILGDCWLLAAIGSLTTCPKLLYRVVPRG QSFKKNYAGIFHFQIWQFGQWVNVVVDDRL PTKNDKLVFVHSTERSEFWSALLEKAYAKLS GSYEALSGGSTMEGLEDFTGGVAQSFQLQRP PQNLLRLLRKAVERSSLMGCSIEVTSDSELES MTDKMLVRGHAYSVTGLQDVHYRGKMETLI RVRNPWGRIEWNGAWSDSAREWEEVASDIQ MQLLHKTEDGEFWMSYQDFLNNFTLLEICNL TPDTLSGDYKSYWHTTFYEGSWRTGSSAGGC RNHPGTFWTNPQFKISLPEGDDPEDDAEGNV VVCTCLVALMQKNWRHARQQGAQLQTIGFV LYAVPKEFQNIQDVHLKKEFFTKYQDHGFSEI FTNSREVSSQLRLPPGEYIIIPSTFEPHRDADFL LRVFTEKHSESWELDEVNYAEQLQEEKVSED DMDQDFLHLFKIVAGEGKEIGVYELQRLLNR MAIKFKSFKTKGFGLDACRCMINLMDKDGSG KLGLLEFKILWKKLKKWMDIFRECDQDHSGT LNSYEMRLVIEKAGIKLNNKVMQVLVARYA DDDLIIDFDSFISCFLRLKTMFTFFLTMDPKNT GHICLSLEQVLGEGWEGICRLAPACPSTPPPPS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SDVPGPASCPRLFPPWDLLPVSTVAADDHVGI EAL
1038	2388	A	8621	3	1494	RSRMARAPLGVLLLLGLLGRGVGKNEELRLY HHLFNNYDPGSRPVREPEDTVTISLKVTLTNL ISLNEKEETLTTSVWIGIDWQDYRLNYSKDDF GGIETLRVPSELVWLPEIVLENNIDGQFGVAY DANVLVYEGGSVTWLPPAIYRSVCAVEVTYF PFDWQNCSLIFRSQTYNAEEVEFTFAVDNDG KTINKIDIDTEAYTENGEWAIDFCPGVIRRHH GGATDGPGETDVIYSLIIRRKPLFYVINIIVPCV LISGLVLLAYFLPAQAGGQKCTVSINVLLAQT VFLFLIAQKIPETSLSVPLLGRFLIFVMVVATLI VMNCVIVLNVSQRTPTTHAMSPRLRHVLLEL LPRLLGSPPPPEAPRAASPPRRASSVGLLLRAE ELILKKPRSELVFEGQRHRQGTWTAAFCQSL GAAAPEVRCCVDAVNFVAESTRDQEATGEE VSDWVRMGNALDNICFWAALVLFSVGSSLIF LGAYFNRVPDLPYAPCIQP
1039	2389	A	8636	1	900	PGRERPGGGGARRRPQHLPALLPSERPDCATL QAMENELPVPHTSSSACATSSTSGASSSSGCN NSSSGGSGRPTGPQISVYSGIPDRQTVQVIQQ ALHRQPSTAAQYLQQMYAAQQQHLMLQTA ALQQQHLSSAQLQSLAAVQQASLVSNRQGST SGSNVSAQAPAQSSSINLAASPAAAQLLNRA QSVNSAAASGIAQQAVLLGNTSSPALTASQA QMYLRAQMLIFTPTATVATVQPELGTGSPAR PPTPAQVQNLTLRTQQTPAAAASGPTPTQPVL PSLALKPTPGGSQPLPTPA
1040	2390	A	8645	98	1388	ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA LTKPKNNHNCKKVSNEEKPKVAIGEECRADE QAFLVALYKYMKERKTPIERIPYLGFKQINLW TMFQAAQKLGGYETITARRQWKHIYDELGG NPGSTSAATCTRRHYERLILPYERFIKGEEDKP LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA PLAPEKDSALVPGASKQPLTSPSALVDSKQES KLCCFTESPESEPQEASFPRLPHHTGHRWQTR MRRRMTNCPPWQITLPTAP
1041	2391	A	8646	113	1492	LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYP GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP NTSLAFVPGVGIKALTNHGTANISTDWGFESP LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF LLISTDLKYETSSKQQPSFHVWEGLNLISRQW RGKSAP
1042	2392	Α	8672	538	170	ARRIARTRESKAAVSQDNVPALQPGKKKKLR LGGKKKKFKFFRLPKEFKKQLMYSPSNFKKM

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion TSLAGNTVQCLNKLKYVIYSAQYPAYGNITT LDMITSTDHVLEQDFWICFTFYSVKERQI GLKTRAPATPTFQREVLGPAKQDMQRRCPRI GLMTSLLKPIKRRWRDYKRWKSGGFTGESC
						HHADTLGDRGGLQGDHSELLQWQKRILRTE GEPSPKYISKNIFPICSYITGFL
1044	2394	A	8718	292	1490	GTVKTSVATPITAGHSCSSGGVLQVKSPATQS GFKFTSKMEDFNMESDSFEDFWKGEDLSNYS YSSTLPPFLLDAAPCEPESLEINKYFVVIIYAL VFLLSLLGNSLVMLVILYSRVGRSVTDVYLL NLALADLLFALTLPIWAASKVNGWIFGTFLC KVVSLLKEVNFYSGILLLACISVDRYLAIVHA TRTLTQKRYLVKFICLSIWGLSLLLALPVLLFR RTVYSSNVSPACYEDMGNNTANWRMLLRIL PQSFGFIVPLLIMLFCYGFTLRTLFKAHMGQK HRAMRVIFAVVLIFLLCWLPYNLVLLADTLM RTQVIQETCERRNHIDRALDATEILGILHSCLN PLIYAFIGQKFRHGLLKILAIHGLISKDSLPKDS RPSFVGSSSGHTSTTL
1045	2395	A	8724	254	3184	FRANLAITVANRRGAQGKMHTCCPPVTLEQ DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY YGEICDNACPCEEKDGILTVSCENRGIISLSEIS PPRFPIYHLLLSGNLLNRLYPNEFVNYTGASIL HLGSNVIQDIETGAFHGLRGLRRLHLNNNKL ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF GKLHLLQVLILNDNLLSSLPNNLFRFVPLTHL DLRGNRLKLLPYVGLLQHMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCETP FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF
1046	2396	A	8736	28	452	SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA
1047	2397	A	8741	673	924	QLIQGGRLIKHEMTKTASA ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL
1048	2398	A	8747	3	5054	PEVTKPSLSQPTAASPIGSSPSPPVNGGNNAKR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq-uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /-possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide deletion, \=possible nucleotide insertion VAVPNGQPPSAARYMPREVPPRFRCQQDHK VLLKRGQPPPPSCMLLGGGAGPPPCTAPGAN PNNAQVTGALLQSESGTAPDSTLGGAAASNY ANSTWGSGASSNNGTSPNPHHWDKVIVDGS DMEEWPCIASKDTESSSENTTDNNSASNPGSE KSTLPGSTTSNKGKGSQCQSASSGNECNLGV WKSDPKAKSVQSSNSTTENNNGLGNWRNVS GQDRIGPGSGFSNFNPNSNPSAWPALVQEGTS RKGALETDNSNSSAQVSTVGQTSREQQSKME NAGVNFVVSGREQAQHINTDGPKNGNTNSL NLSSPNPMENKGMPFGMGLGNTSRSTDAPSQ STGDRKTGSVGSWGAARGPSGTDTVSGQSNS GNNGNNGKEREDSWKGASVQKSTGSKNDS WDNNNRSTGGSWNFGPQDSNDNKWGEGNK MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGHPLLENQGNAQAPCWGRSS STGSEVEGQSTGSNHKAGSSDSHNSGRRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATQTKNSGGWGDAPSQSNQMKSGWGELS AATEWKDPKNTGGWNDYKNNNSSNWGGGR PDEKTPSSWNENPSKDQGWGGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWG EGGQNEIGTWGNGSMSKMQQQPPQQPPPPP PEASGSWGGPPPPPPGNVRNSNSWSGPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGRAPREPNLPTP MTSKSASDSKSMQDGWGSSDGPVTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWGQG GKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVGRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GYYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPOLSPQQQQLQQRQPGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGWVGGKEAGTESRFKQWTSMME GLPSVATQEQQQQLQLQRQRGMKHSPSHPVGPK FHLDNMVPNALNVGLPPLQTKGPIPGYGSGF SSGGMDYGWVGGKEAGTESRFKQWTSMME GLPSVATQEARMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK HISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLATICMQHGLLTFHLNLTQGTA
						RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS
1049	2399	A	8748	200	1387	LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF
						LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA PFALSALLYGANNNLVIYLQRYMDPSTYQVL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				·		SNLKIGSTAVLYCLCLRHRLSVRQGLALLLL MAAGACYAAGGLQVPGNTLPSPPPAAAASP MPLHITPLGLLLLILYCLISGLSSVYTELLMKR QRLPLALQNLFLYTFGVLLNLGLHAGGGSGP GLLEGFSGWAALVVLSQALNGLLMSAVMKH
						GSSITRLFVVSCSLVVNAVLSAVLLRLQLTAA FFLATLLIGLAMRLYYGSR
1050	2400	A	8758	3	1660	WVSSMGFEELLEQVGGFGPFQLRNVALLALP RVLLPLHFLLPIFLAAVPAHRCALPGAPANFS HQDVWLEAHLPREPDGTLSSCLRFAYPQALP
				_		NTTLGEERQSRGELEDEPATVPCSQGWEYDH SEFSSTIATESQWDLVCEQKGLNRAASTFFFA GVLVGAVAFGYLSDRFGRRRLLLVAYVSTLV LGLASAASVSYVMFAITRTLTGSALAGFTIIV
		,		·		MPLELEWLDVEHRTVAGVLSSTFWTGGVML LALVGYLIRDWRWLLLAVTLPCAPGILSLWW VPESARWLLTQGHVKEAHRYLLHCARLNGR
						PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA
						GRRLTQAGTLLGTALAFGTRLLVSSDMKSWS TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS
						LPKLTYGGIALLAAGTALLLPETRQAQLPETI QDVERKSAPTSLQEEEMPMKQVQN
1051	2401	A	8759	515	1625	EIRTPVÄVSSAPSGDSEGDEEETTQDEVSSHTS EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV
						YQHTAAVVSAKSYMCPVCGRALSSPGSLGR HLLIHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP
						AGILLVCNNCAAYRKLLEAQTPSVRKWALRR QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR
						EAMRLKRANETPEKRQARLIREREAKRLKRR LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH
1052	2402	A.	8763	1106	70	RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY
					٠	PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP
						RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL
					ļ	LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY
1053	2403	A	8768	2	712	RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW
						PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK
						VNEILNRLEKTKVERFPDLAAEKECRDREER NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM
1054	2404	Α	8769	344	527	REATTLACRNSCWVFSRCSLGACKPTVCSMP SLSRQGSQTLCLRLAEYCMESVDSQRLLLS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ł	{	ł		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1			ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
ŀ				sequence		nucleotide insertion
1055	2405	A	8770	430	1104	QQESPAAGAARMNCKEGTDSSCGCRGNDEK
						KMLKCVVVGDGAVGKTCLLMSYANDAFPEE
	İ		1			YVPTVFDHYAVTVTVGGKQHLLGLYDTAGQ
		ł			i	EDYNQLRPLSYPNTDVFLICFSVVNPASYHNV
i			1	[QEEWVPELKDCMPHVPYVLIGTQIDLRDDPK
ŀ	{	1	ĺ	[į	TLARLLYMKEKPLTYEHGVKLAKAIGAQCYL
						ECSALTQKGLKAVFDEAILTIFHPKKKKKRCS
]			1	i	1	EGHSCCSII
1056	2406	A	8773	261	332	NPRIQLSGNSCCAGSCRVWLSEQ
1057	2407	A	8778	3	477	PAGIRHEQARGADRMGKCRGLRTARKLRSH
105.		**) ~		RRDQKWHDKQYKKAHLGTALKANPFGGAS
						HAKGIVLEKVGVEAKQPNSAIRKCVRVQLIK
1				•		NGKKITAFVPNDGCLNFIEENDEVLVAGFGR
1	1		l	,	t	KGHAVGDIPGVRFKVVKVANVSLLALYKGK
1			1	ĺ	Í	KERPRS
1058	2408	A	8808	171	881	PGLSQEPSGSMETVVIVAIGVLATIFLASFAAL
1038	2408	1	0000	1 * ′ *	001	VLVCRQRYCRPRDLLQRYDSKPIVDLIGAME
İ			1			TOSEPSELELDDVVITNPHIEAILENEDWIEDA
Į.		1		1	l	SGLMSHCIAILKICHTLTEKLVAMTMGSGAK
İ		1	İ			MKTSASVSDIIVVAKRISPRVDDVVKSMYPPL
				l		DPKLLDARTTALLLSVSHLVLVTRNACHLTG
1	1			ł		GLDWIDQSLSAAEEHLEVLREAALASEPDKG
i				!		LPGPEGFLQEQSAI
1059	2409	A	8809	246	757	MRLQGAIFVLLPHLGPILVWLFTRDHMSGWC
1039	2409	Ι ^	0009	240	131	EGPRMLSWCPFYKVLLLVQTAIYSVVGYASY
ł	Ì	1		1		LVWKDLGGGLGWPLALPLGLYAVQLTISWT
1	Į.	l	ŀ	İ		VLVLFFTVHNPGLALLHLLLLYGLVVSTALI
1	1	l		ļ.		WHPINKLAALLLLPYLAWLTVTSALTYHLWR
1	1	ł		İ		DSLCPVHQPQPTEKSD
1070	0410	1	0010	304	201	
1060	2410	A	8810	304	381	PKLSVYPLQSHHCLSEPFQSLVCCLA
1061	2411	Α	8820	1673	848	SCKTENLLEMWWFQQGLSFLPSALVIWTSAA
ĺ	1	1		[(FIFSYITAVTLHHIDPALPYISDTGTVAPEKCLF
	i					GAMLNIAAVLCIATIYVRYKQVHALSPEENVI
1		1				IKLNKAGLVLGILSCLGLSIVANFQKTTLFAA
			1			HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH
		1	1	1		GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS
			I	1	1	GNFGTDLEQKLHWNPEDKGYVLHMITTAAE
1	}		1	1		WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL
	1	 	1	<u> </u>	1.50	TLYDTAPCPINNERTRLLSRDI
1062	2412	A	8824	1	763	GGAPPASVPARESPVSGAQGSSRTRGHKRAA
	1		1	1		GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA
1		1		1		YSLAPATPEVKVACSEDVDLPCTAPWDPQVP
	1	1		l	1	YTVSWVKLLEGGEERMETPQEDHLRGQHYH
		1	1	1		QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR
1	1	l		ł	<u> </u>	CTLQDPDGQRNLSGKVILRVTGCPAQRKEET
1	1	1				FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI
		1				FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT
	1	1				ELV
1063	2413	A'	8826	147	627	CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE
1						HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT
		1			1	AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ
		1			1	KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH
		1				CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL
}	1	1		}	1	TR
1064	2414	Ā	8835	2982	1869	LKDTLKSQMTQEASDEAEDMKEAMNRMIDE
	~	1	1	1		LNKQVSELSQLYKEAQAELEDYRKRKSLEDV
1	1	1				TAEYIHKAEHEKLMQLTNVSRAKAEDALSE
	ŀ				ļ	MKSQYSKVLNELTQLKQLVDAQKENSVSITE
	[1				HLOVITTLRTAAKEMEEKISNLKEHLASKEVE
<u> </u>	1			<u> </u>		1 Z

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	}	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	•	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		Ì		residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	/=possible nucleotide deletion, \=possible
1			i	sequence		nucleotide insertion
						VAKLEKQLLEEKAAMTDAMVPRSSYEKLQS
		ĺ				SLESEVSVLASKLKESVKEKEKVHSEVVQIRS
						EVSQVKREKENIQTLLKSKEQEVNELLQKFQ
						QAQEELAEMKRYSESSSKLEEDKDKKINEMS
					i	KEVTKLKEALNSLSQLSYSTSSSKRQSQQLEA
	1				ļ	LQQQVKQLQNQLAECKKQHQEVISVYRMHL
ļ)	LYAVQGQMDEDVQKVLKQILTMCKNQSQK
1065	2415	A	8841	7	662	K
1065	2413	A	8841	3	663	AAATAASLSPRGCRLRTPSSDVGPSRAPPPSA
]			APLPTGRAQMSPSGRLCLLTIVGLILPTRGQTL KDTTSSSSADATIMDIQVPTRAPDAVYTELOP
1		İ	i i	i		TSPTPTWPADETPQPQTQTQQLEGTDGPLVT
j						DPETHKSTKAAHPTDDTTTLSERPSPSTDVQT
1					ļ	DPQTLKPSGFHEDDPFFYDEHTLRKRGLLVA
	L					AVLFITGIILTSGKCRQLSRLCRNHCR
1066	2416	A	8853	3806	2204	FVGEQEGGCEAGAGRGAQTYPGEAGERWFG
1]]			ļ	RRRRRGRVVSRKKMSLKSERRGIHVDQSDLL
	-					CKKGCGYYGNPAWQGFCSKCWREEYHKAR
						QKQIQEDWELAERLQREEEEAFASSQSSQGA
						QSLTFSKFEEKKTNEKTRKVTTVKKFFSASSR
		}				VGSKKEIQEAKAPSPSINRQTSIETDRVSKEFIE FLKTFHKTGQEIYKQTKLFLEGMHYKRDLSIE
,		1				EQSECAQDFYHNVAERMQTRGKVPPERVEKI
						MDQIEKYIMTRLYKYVFCPETTDDEKKDLAI
						QKRIRALRWVTPQMLCVPVNEDIPEVSDMVV
					[KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI
	1					KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI
		1				QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE
					}	KLDAQSLNLSQEDFDRYMSGQTSPRKQEAES
1						WSPDACLGVKQMYKNLDLLSQLNERQERIM
1					;	NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI
1067	2417	A	8855	1372	1513	KPPNQPLAAIDSENVENDKLPPPLQPQVYAG SNMREVGCGWLVPVIPAFWEAEVGGSLEARS
1007	2417	^	0000	1372	1313	LRQAWATKODPISKKK
1068	2418	A	8856	1530	1583	PCRPGMECNSMISVHCNL
1069	2419	A	8857	1530	1583	PCRPGMECNSMISVHCNL
1070	2420	Α	8866	293	1675	PYPQGGYPQGPYPQGGYPQGP
						YPQSPFPPNPYGQPQVFPGQDPDSPQHGNYQ
		}]			EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF
		1	[LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV
		1				WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL
						VALSVLTASLSYMVGMIASFYNTEAVIMAVG
						ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM
		l				VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA
1	l	1				VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP
						WHGSASCTSPLSCPQAQPREKDASLQPSCMY
		1				TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC
			[HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ
1	1	ĺ				EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS
						GDMRSGGLIPVLSPE
1071	2421	A	8868	2	358	ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH
	}		J i			DDKMGSNTFFKRNDCRYVMISCKADMAYDN
		1				VRHPFMI*SI\KLIMEETYLNIIKAVYDRPTASII
1005	10400	ļ.,	00.00			LNGEKLKVFPVRSGT*QGCSVWP
1072	2422	A	8870	33	658	MESVLSKYEDQITIFTDYLEEYPDTDELVWIL
1	1	· ·				GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH
		1	!			LGRDWSWEKQKEQPKEYQRILOCFLDRKDC
L	L			L.—	L	LOUVE ARE A STATE A STATE OF THE PARTY OF TH

CECTO	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	liou	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	uence		09/496		to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	i	914	correspondi	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence	1	ł	914	ng to first amino acid		T=Threonine, V=Valine, W=Tryptophan,
}				residue of	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1				sequence	
1	[1	peptide	1	/=possible nucleotide deletion, \=possible
	<u> </u>		ļ	sequence		nucleotide insertion
						CYSIHQMAQMGVGEGKSIGEWVLGPNTV\AQ
	ł					GV*KNLA\LFDEW\NSLGLVYVSM\DNPSGSIA
	<u> </u>		<u> </u>			RFPKKLCRVLPL\SADTAGLTGP
1073	2423	Α	8879	146	412	DFSV*GDVDIEVTCPICLQLLTEPLSLNCGLRL
		1		i		*QVCITA*IKESVIISGG*SSSPVCHTTFQPANL
			L			RTSRYLPT*SIKSLGPDEPQEG
1074	2424	A	8884	67	435	HLQGRSIRTLQLTGENEKNCEVSERIRRSGPW
1	ļ	į				KEISFGDYICHTFQGDCWADRSPLHEAAAHG
	ļ.		ļ			RLLALKTLIAQGVNVNLWTL/DRVSSLHEACL
		1				*GPVACAKPYWKMVPRHGGTVTGPPLLMV
1075	2425	Α	8896	1294	248	RSGDRNGLTHQLGGLSQGSRNQSYRSRSRSR
						SRERPSAPRGIPFASASSSVYYGSYSRPYGSDK
•		}	1	}	}	PWPSLLDKEREESLRQKRLSERERIGELGAPE
1	1					VWGLSPKNPEPDSDEHTPVEDEEPKKSTTSAS
i	1					TSEEEKKKKSSRSKERSKKRRKKKSSKRKHK
						KYSEDSDSDSDSETDSSDEDNKRRAKKAKKK
ļ						EKKKKHRSKKYKKKRSKKSRKESSDSSSKES
i	i					QEEFLENPWKDRTKAEEPSDLIGPEAPKTLTS
		ļ	ŀ			QDDKPLNYGHALLPGEGAAMAEYVKAGKRI
l		1	Į	}	1	PRRGEIGLTR*RNCHHLNAQVM**VVSRHRR
	ļ		İ			MEAVRTAKREPESTVLMRREPLHPFNPRRET
	1					KERE
1076	2426	A	8899	146	789	GRSTEAEKEPAFDERTGKGRRLPRAGEFHG*E
1		}				*APGPGPRSFQVSRKMPEE\PPGARKHPFSGKS
ļ	ł			į		FYLDLPAGKNLQFLTGAIQQLGGVIEGFLSKE
į		l				VSYIVSSRREVKAESSGKSHRGCPSPSPSEVR
ì	ł	1	ļ		i	VETSAMVDPKGSHPRPSRKPVDSVPLSRGKE
	1		1			LLQKAIRNQK**CTVQQLSHCRLY\GEKTTAK
i					į	RSQREHVQQQSQEHGKWPDLKGPR
1077	2427	A	8901	352	3	AKIGAYKYIQELWRKKQSDVMHFLLRVRCW
1	} = '-'	1	0,0.		1	QYPALHRAGTEWQLSALHRAPRSTQPDKAC
ŀ	1	1			1	RLGYKAKQGYIIYRICVRRGGWKCPVPKAVT
ļ	1	1	ļ			\YGKPVHHGVN*LKFAQSLQSVAEEQ
1078	2428	A	8905	536	781	ACPAENREVPEMAAGQAPHAGPGAGPGQPA
1070	2420	^	0,00	330	/01	PALPFAATPGSRGQALCRGGRRRQHLHGPLH
			l	1		RP*QAAPALHAGCQLAPHPPT
1070	2420	Ι	8012	121	376 .	The sales are an area of the sales are also and the sales are are also are
1079	2429	A	8912	121	376 ·	NLIWKLCVTERRLVILDNYDLASE/YEANKYI CNRIIQFKPGQDKYFTLGLPTGSTPL*CYPKLI
		l				EYNKNGHLSFKYVKTFSMDEY
1080	2430	 _ _ _ 	8920	381	1788	SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG
1000	2430	Α	0720	301	1/00	
1		Ι,	1	Ī		GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY
	•	}				DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL
J		J	ļ	ļ		YAPICMEYGRVTLPCRRLCQRAYSECSKLME
]					MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA
1	}	l	1			GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY
]	1	l	1	}		SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS
1]	IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV
	1	1	1			WHMMVSLIFF\IGFLLEDRVACNA\SIPAQYKA
	1	ì	1			STVTQGSHNKACTMLFMILYFFTMAGSVWW
1	1	1	}	ļ		VILTITWFLAAVPKWGSEAIEKKALLFHASA
[[l				WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD
						VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR
[1	l	ł	ĺ	1	VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL
<u></u>		l				VVIGCYFYEQAYRGIWETTWIQERC
1081	2431	A	8922	56	420	EERTKMSTGPDVKATVGDISSDGNLNVAQEE
-					1	CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG
1	}	1	-	}		TLANFVF\CSVRHGLALILQLCNFSIYTQQMN
l			1			LSIAIPAMVNNTAPPSQPNASTERPST
1082	2432	A	8923	355	1079	PFGTPSSTMAVVKNKCLMKGGKKGVKKKVV
				L		

muclecide seq	SEQ ID	SEQ ID	Met hod	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
USSN Osadion	NO: of	NO: of	noa	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
			1				I=Isoleucine K=I voine I =I evoine
1083							M=Methionine N=Asparagine P=Proline
aninto acid residue of peptide residue of peptide residue of peptide sequence T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosis, X=Unknow, Y=Siop codon, V=possible nucleotide deletion, V=possible nucleotide deletion, V=possible nucleotide insertion T=Threonine, V=Valine, V=Tyrosis, X=Unknow, Y=Siop codon, V=possible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Valine, V=Tyrosible nucleotide insertion T=Threonine, V=Tyrosible nucleotide insertion T=Threonine, V=Tyrosible nucleotide insertion T=Threonine, V=Tyrosible nucleotide insertion T=Tyrosible nucleotide insertion T=Tyrosible nucleotide insertion T=Tyrosible nucleotide insertion T=Tyrosible nucleotide insertion T=Tyrosible nucleotide insertion T=Tyrosible nucleotide insertion T=Tyrosible nucleotide insertion T=Tyrosible nucleotide insertion T=Tyrosible nucleotide insertion T=Tyrosible nucleotide insertion T=Tyrosible nucleotide T=Tyrosible nucleotide insertion			1				O=Glutamine R=Arginine S=Serine
Persidue of peptide sequence Y=Tyrosine, X=Uluknown, *=Siop codon, / Possible nucleotide tellon, Possible nucleotide tinosertion GPFSKKDOYDVKAPAMFNIRNTGK/TLVART QGTQIASDGLKGLIFEVSLADLQNDEVAFRK REALTED/QOENCLINFY GMD TLOBALCSMV EKWSTMIEAH7DVKTTDGYFFHLECVGFTKK HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLKETSY*** HNNQLGSTAMSCLILLIDLITLLILLIMLIGH AGYSGQLAGVAVSAGSPURYKFHVEPYGET GWLLTESCSISPKLICASY** GWLLTESCSISPKLICASY** HNNQLGSTAMSCLILLIDLITLLILLIMLIGH AGYSGQLAGVAVSAGSPURYKFHVEPYGET GWLLTESCSISPKLICASY** GWLLTESCSISPKLICASY** HNNQLGSTAMSCLINFY** HNNQLGSTAMSCLINFY** HNNQLGSTAMSCLINFY** HNNQLGSTAMSCLINF** HNNQLG							
peptide							Y=Tyrosine, X=Unknown, *=Stop codon.
		ŀ			peptide	•	
QCTQIASDGILKGILFEVSLADLQNDEVAFRK FRLITEDYQOENCL INFYGMDLICOMECISM ERWSTMIEAHYDVKTTDGYFFHLFCVGFTKK HNNQLKISTAY-HQOSQIQKKMMEMT'EV OTNDLKEVYNKLIPDNIGKDTEK-VCPIYPLH DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMER DVFIRK-VKMLENDPR		ĺ	{	Ï			
QCTQIASDGILKGILFEVSLADLQNDEVAFRK FRLITEDYQOENCL INFYGMDLICOMECISM ERWSTMIEAHYDVKTTDGYFFHLFCVGFTKK HNNQLKISTAY-HQOSQIQKKMMEMT'EV OTNDLKEVYNKLIPDNIGKDTEK-VCPIYPLH DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMELRGOGOSS DVFIRK-VKMLENDPREPMER DVFIRK-VKMLENDPR					i		GPFSKKDQYDVKAPAMFNIRNTGK/TLVART
EKWSTMEAH7DVKTTDGYFFH,FCVGFTKP, EVGFTYPH		j					QGTQIASDGLKGLLFEVSLADLQNDEVAFRK
HNNQLKITSYA+HQQSRQIQKKMMEMITYE							
OTNDLKEV/NRLIPDNIGKDTEK/MCPITYPLE							
DVFIRKVKMLENPGFERNERGGGSSS							OTNDLKEVVNKLIPDNIGKDTEKV/CPIYPI.H
1083	1			[
WGAVQGRAMSDLILLIDTLLILIM_LIGF	1083	2433	A	8948	28	385	
1084 2434 A 8950 156 318	'		'				
1084	ļ			ļ			AGYSGQLAGVAVSAGSPPI/RYKFHVEPYGET
1085 2435 A 8956 16 413 HMGQLGYFIQCWBCKRLISFWKTI*QSPAK							
1085	1084	2434	A	8950	156	318	HYTPINTDTIENSENNKCW*GY*E\VGLIHHW
*TIYTSYDTAPISIGI/YPKRMSSKCHOETCAR MFILAPFTATIKGKQLTCPLVEERIDY\MWYS HKYTIKVKRNL*VTITHTWV\NLNILMFEILLW YSHKYY YSHKYY YSHKYY YSHKYY YSHKYY YSHKYY YSHKYY YSHKYY YSHKYY YSHKYY YSHKYY YSHKYY YSHKYY YSHKYY YSHKYY YSHKYY YSHGGESEARLHRCTPAWIT THYRILIGHED_UTSYCHCHILMPVS*ELQRL *ERSVCAFHVCIQITYVCLQVYACMCVYYICM FVYSVYGGGLCTCVCMDVYICVCVQEFL YSVYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYSYYGGGLCTCVCMDVYICVCVQEFL YSWYGGGLCTCVCMDVYICVCVQEFL YSWYSYYGGGLCTCVCMDVYICVCVQEFL YSWYSYYGGGCGCTCVCMDVYICVCVQEFL YSWYSYYGGGCGGTCTLSCDMGVV YSWISTEPSSLEDLEATGTGTSKTPLSDMGVV YSWISTEPSSLEDLEATGTGLSDMGVV YSWISTEPSSLEDLEATGTGLSDMGVV YSWISTEPSSLEDLEATGTGLSDMGVV YSWISTEPSSLEDLEATGTGLSDMGVV YSWISTEPSSLEDLEATGTGLSDMGVV YSWISTEPSSLEDLEATGTGLSDMGVV YSWISTEPSSLEDLEATGTGLSDMGVV YSWISTEPSSLEDLEATGTGLSDMGVV YSWISTEPSSLEDLEATGTGLSDMGVV YSWISTEPSSSPFPDHFDHFLFT YSWISTEPSSSPFPDHFDHFLFT YSWISTEPSSSPFPDHFDHFLFT YSWISTEPSSSPFPDHFDHFLFT YSWISTEPSSSPFPDHFDHFLFT YSWISTEPSSSPFPDHFDHFLFT YSWISTEPSSSPFPDHFDHFLFT YSWISTEPSSSPFPDHFDHFT YSWISTEPSSSPFPDHFDHFT YSWISTEPSSSPFPDHFDHFT YSWISTEPSSSPFPDHFDHFT YSWISTEPSSSPFPDHFDHFT YSWISTEPSSSPFPDHFDHFT YSWISTEPSSSPFPDHFDHFT YSWISTEPSSSPFPDHFDHFT YSWISTEPSSSPFPDHFDHFT YSWISTEPSSSPFPDHFDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT YSWISTEPSSSPFPDHFT							
MFILAPFTATIKGKQLTCPLVEREIDYMWYS	1085	2435	Α	8956	16	413	HMGQLGYFIQCWWECKRLISF\WKTI*QSPAK
HKYYKKRRL*VTITHYWVNLNILMFEILW YSHKYY YSHKYY 1086							
1086			1				MFILAPFTATIKGKQLTCPLVEERIDY\MWYS
1086		1	1	ļ			
1087	1006	2426	 	00/2	-	1006	
1087	1086	2430	A	8962	808	1026	
*ERSVCAFHVCIQTYVCLQVYACMCVYYICM FYYSVYGCILCTCVCMDVTICVCVQEFL	1097	2427	_	0005	50	320	
FYYSVYGGLCTCVCMDYYICVCVQEFL	1007	2431	^	0703	36	330	
1088	1	l	ł	1	}		FVVVVCCI CTCVCMDVVICVCVCET
KKFLDMMSNA*STKKHDKLD/LIKFKT/LCSA KYTVKRIKHRPTDLEKMLRNHLSDKD*YS/GV YKDLSKLNRKTE/S*/VKKWVKDLSRYFIKE VISMENKHKKIFSTS VISMENKHKKIFSTS VISMENKHKKIFSTS VISMENKHKKIFSTS VISMENKHKKIFSTS VISMENKHKKIFSTS VISMENKHKKIFSTS VISMENKHKKIFSTS VISMENKHKKIFSTS VISMENKHKKIFSTS VISMENKHKKIFSTS VISMENKHKKIFSTS VISMENKHKKIFSTS VISMENKH V	1088	2438	Δ	8080	304	404	
KYTVKRIKIHPTDLEKMLRNHLSDKD*YS/GV YKDLSKLNRRKTE/\$*/VKKWVKDLSRYFIKE VISMENKHKIFS*1/S** 1089	1.000	2130	^	6767	354	404	
YKDLSKLNRRKTES* / VKKWVKDLSRYFIKE							
VISMENKHKKIFSTS							
NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV GVEDNAYTLEVNSRYMRAVGIM*IHL 1090 2440 A 8996 2 351 SNITITLT*MKKYDNTFCW*GCQIG/7LIYC WCESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI 1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQVCKDS FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHIPPHIPV*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQFAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFDGMVEQDKRDEWVKRR LKNNREISRFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEIL.PL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVOTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEPRQEVPMCTDSEP							
NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV GVEDNAYTLEVNSRYMRAVGIM*HHL	1089	2439	Α	8991	60	329	MALTPESPSSFPGLAATGSSVPEPPGGPNATL
1090 2440 A 8996 2 351 SNITITLT*MKKYDNTFCW*GCQIG/T/LIYC WQESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYMKMLHTR LLLAALFIIVQYWKQSKDHYI 1091 2441 A 8997 97 456 YPLPVC\$YL\$GPRGEHWN\$LGGKS\$CPLPLPT LVSSRFKISKVIVVGDL\$VGKTCLINR*GGAG AELGRVGP\$LARWAGSR\$CHLVP\$QIVCKD\$ FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SFIKRHILIFEDDWHQTTCCHIPHHPFF*RCQ FHIFYV\$VQNSI\$P\$L\$V\$SS\$HPDRPDHEVHQH RAAHHHQHGQPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQ\$LIL\$R\$SVMGVRGLQGFVG\$ TCPHICTVVNFKELAEHHR\$KYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYF\$ALRDF VKTFTAAGIKLIFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIP\$GLA VFTRFALKTLGQETLC\$LQEADYEVA\$YGLQ HNCLGILGEDTDYLJYDTCPYF\$ISELCLE\$LD TVMLCREKLCE\$LGLVADLPLACLLGNDII PEGMFE\$FRYKCL\$YTYVENFDKKGNIILA V\$DHISKVLYLYQGEKKLEEILPLVTKQ\$\$FL *RNGISFTRT/INLHGF\$KNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/P\$PRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTD\$EPRQEVPMCTG\$PEPRQEVP MCTGPEPRQEVP MCTGPEPRQEVPMCTD\$EPRQEV							NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV
WQESKFIQAFWSKIQQYLA*ISHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLLAALFIIVQYWKQSKDHYI 1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS FDKNYKAPJGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPVF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLJYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEELIPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTINK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPMCTDSPPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSE			l				
GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI 1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPVF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTINK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHII/PERQEVPTCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP	1090	2440	A	8996	2	351	
LLIAALFIIVQYWKQSKDHYI				В			
1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQNVCKDS FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPF*RCQ FHIFYVSVQNSISPSLSVSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFTDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSPPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGSEPRQEVP	1	}					
LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHIPPHPPF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHIQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP	1001	0.00		2005		455	
AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS FDKNYKAPIGADFEMERFEVLGIPF 1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPVF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMPESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP	ועטנ	2441	A	8997	9/	436	* * * * * * * * * * * * * * * * * * *
FDKNYKAPIGADFEMERFEVLGIPF				1			
1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPVF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY]	ļ] .			
FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHIHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMYTGPESRQEVPMY	1092	2442	Α	8000	548	811	
RAAHHHQHGQGPLGHGLVARVG 1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCTGDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY		2772		0,,,	770	011	
A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNILLA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY				1	İ		
TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMYT	1093	2443	A	9002	3	2745	
AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMY						·	
VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMYT				1			
LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMYTGPESRQEVPMY			1				
VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	}	}	l	1 .			
TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY				!			VFTRFALKTLGQETLCSLQEADYEVASYGLQ
PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							HNCLGILGEDTDYLIYDTCPYFSISELCLESLD
VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							
*RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							
RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL
GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							
PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY				İ			GEVERCILLIBERROEURTCORRERACIONES
MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY			ĺ	[DESERVENCE LYCENTICS DEFENDACE DE COMP
RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							MCTGPEAROEVPMCTDSEPROEVPMCTDSEP
	1			, 1	1		TGPESRQEVLIRTDPESRQEIMCTGHESKQEV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid	Predicted end nucleotide location corresponding to last amino acid residue of peptide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						PICTDPISKQEDSMCTHAEINQKLPVATDFEFK LEALMCTNPEIKQEDPTNVGPEVKQQVTMVS DTEILKVARTHHVQAESYLVYNIMSSGEIECS NTLEDELDQALPSQAFIYRPIRQRVYSLLLED CQDVTSTCLAVKEWFVYPGNPLRHPDLVRPL QMTIPGGTPSLKILWLNQEPEIQVRRLDTLLA
						CFNLSSSREELQAVESPFQALCCLLIYLFVQV DTLCLEDLHAFIAQALCLQGKSTSQLVNLQP DYINPRAVQLGSLLVRGLTTLVLVNSACGFP WKTSDFMPWNVFDGKLFHQKYLQSEKGYA VEVL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT HHWPSPLGLTPRREVGKTGLQLPQDGLWV
1094	2444	A	9021	97	834	AREACRAKTDFPGRRFRLWPSCCCRVIVGAE T*H\MAEPVSPLKHFVLAKKAITAIFDQLLEFV TEGSHFVEATYKNPELDRIATEDDLVEMQGY KDKLSIIGEVLSRRHMKVAFFGRTSSGKSSVI NAMLWDKVLPSGIGHITNCFLSVEGTDGDKA YLMTEGSDEKKSVKTVNQLAHALHMDKDLK AGCLVRVFWPKAKCALLRDDLVLVDGPGTD VTTELDSWIDKFCTKSSTREITNSGSDT
1095	2445	A	9022	1	537	LVLNSRVEDFVPPEGAGRTLPFALRPLAACW LLHRRARRSSALCPRPRSWGVSGGEGAGARE P*ITSSSCCLSAA/SHLSIQSPNMAGARRRIRPQ LAKEKIEGCHICTSVTPGEPQVFLGKDKAFTF DYVFDIDSQQEQIYIQCIEKLIEGCFEGYNATV FAYGOT\GAGKTYTMGTGFD
1096	2446	A	9029	1	285	FFFFNVCKSPKVPKPGCKEESTGTLFKNTLISL GQHSETPSLKKK\LAGYSGMCL*SQVLRRLRQ EDCLSPGGGNCRES*SCPYTPAWITERDPV
1097	2447	A	9032	716	357	ARSTGFWGEILWCGFLKRSLALSPRVKCSGAI LAHCNFRHAGFPPLSCLSLPNRWEYRRPPARP GKFFLVFLVETGFQC/G*DGLDLLTSRSACLG LPKCWDYRREPAASIIFQTTFFINSK
1098	2448	A	9038	230	652	KVVVMSCEDINISGSFYRNKLKYLAFLCKRTS TNPSQGPYHLWVPSHIFWQTTCGRLPHKTKQ G*AALDHLKVFDRIPLPYDKKKQMAVSATLE VVRPKP*RKFAYLGHWAQKVDWKYQAMTA TMGEKRKVYYQKICYQKK
1099	2449	A	9043	185	372	IIFYSHQQCMRV/WQGCGDIETLIHCW*E*KII HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQKWKKINI
1100	2450	A	9045	763	584	ROSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF
1101	2451	A	9050	2,75	2	LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS
1102	2452	A	9053	449	1224	KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV
1103	2453	A	9058	403	3	GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW*GC\GA AGMQI/H/CW\WCVNVGKFWEMS*YYLLKLSI ST/PYDPAIPLLGIYL*ETRVYIHPKTCMRMLIA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion APFVLAVNC
1104	2454	A	9064	75	393	KWLFSSLNITGRGDIIGHLKWLDCR\NCSSFPI KRNRQTHSTESNKLKAGHSFGYN*LIH*NS\V KTDCGCGANSKGVVVVMKV\KTAQQKQTTS YMQIGTTKNSRAT
1105	2455	A	9065	366	778	DLLILRNLAFPELKRRNCISRFYLAYHLHKIYS RSILLCNNCSGFYILSL*QYDVFFFNYFFFRDR AWPCCPGWSAAWLTIVILAHYRRPGLERSCC LSLSSSWDHRRVPPCPANF*/YFSMGFTAFPRL VLNS*TQGI
1106	2456	Α	9083	673	816	ESGSLIH*WWENKPAQPLWWEI*QHVQKLPT HFPCDPAIPLLGICPED
1107	2457	A	9086	580	18	KPSSGSFIRAIYIFLSTAHVPALFSVLVRTKLT* AFSQSSVLWAHKQQKTSLSLVIR/ERLQIKTA VRENFLPIRLAKILKLDNVKCWQG/SGSNMSL I/HCWWEYNVIHIIWNSVTFPRKVEHVYITYA PEISVR*IHGGLPTLVHQETHTSVFRGAPSVIP ETR\CRPTKESINKLLHIYTMEHYGDENK
1108	2458	Α	9093	540	1	GGNDCSVTPTTEPGRKEIT*KRKF*EKTDRLP GA/PPSRTPPTPYPCPHGDRLLPPSRPLPAGPA SAFPPAERSRGHRRASL*RARWSAAVPRRSA GSASEPVQSRWLRLPVGSDSPPAVPVRVCPAP DSRPAAPGSRLPDPGLDSPAPSRTPSSSVD*GG QRPPPPSGDSLSPPGCCRY
1109	2459	A	9099	1255	1425	HESYHVNPNLCNPVAPTSGAHSIG*KWPSWL GAVAHSCNPSTLVGRGGRITRGOELR
1110	2460	Α	9103	242	70	EEQFFFFAVGMFP*VDFLAPASGELWDRLRLT CSRPFTRHQSFGLAFLRVCSSLDSLDDSVVGP SALLSSVL/NQGGRNVLEAREAAKHPTI*RQS LLRKQRNKRMAIP
1111	2461	Α	9110	189	121	SFLSVRLECNGAIMAHCALPLPG
1112	2462	A	9113	100	910	RRRGGGSRPRRTPVPAPGPGPSFGMDVRFYP AAAGDPASLDFAQCLGYYGYSKFGNNNNYM NMAEANNAFFAASEQTFHTPSLGDEEFEIPPIT PPPESDPALGMPDVLLPFQALSDPLPSQGSEFT PQFPPQSLDLPSITISRNLVEQDGVLHSSGLHM DQSHTQVSQYRQDPSLIMR\PSST*PDAARSG VMPPAQLTTINQSQLSAQLGLNLGGASMPHT SPSPPASKSATPSPSSSINEEDADEANRAIGEK RAAPDSGKKPKTPKK
1113	2463	A	9120	3452	3051	FLRPSFALVPQAGVQWCALSWLQPPSPRFK*F SCLSLPSSWDYRHVPPRPANFFVLLVETGFLH VGQAGHEPLTSGDPPASASQSAGITGVSHQA WPSFFIFSRDTVLLCCSGWSRTSGLKQSACLS LLKCWDY
1114	2464	A	9122	152	377	NQLPLQQWTFFIYETGFCSVAQAGVQCRDHS SLHP*PPG\SSDPPAPPS*VLGITGQRYHACLII YLYVQTVPQRV
1115	2465	A	9124	553	981	QRPLLRQQLGSWPTCRSLEGDLASPW**RLPG SPRMRRSGT/ATLNLPLSPQGTVRTAVEFQVM TQTQSLSFLLGSSASLDCGFSMAPGLDLISVE WRLQHKGRGRGDLHLPDHHLSVPSSADHPA QQPSQFNGRNLYFLPLFR
1116	2466	A	9135	48 .	410	SASHEPAEHDGGADSLSASQPPRPAGRPAGA QHVHVPPWTDVLAGQDRRAPTAGDGAPWP APGGHVPSTRPHDPAEFHADEAAGRGGRGLQ PAAPHALPAGLPHGPPAPA/PAEGGGTP*GSA GAGGP*GSPAGRACGAAGCRPRPPRPAASSA *NSAGS*GLVEGT*PPGAGHGAPSPAVGARLS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion CPARTSVQGGTWTC*APAGRPAGLGGWEAE RESAPPSCSAGS*DAD*GAEPWGAGSRSWGS KSGHWAKECLQPRIPPRPCPICVGPHWKSDCP TCPGAVPRAPGTLPQGSLTDSFPDLLSLVAED *CCLMASEASWTIVELWVTLTVEGKSVP/CL NTEATHSTLPSFQGPVSLASITVVGIDGQASKP LKTPQLWCQLGQYSFMHYFLVIPTCPVPLLG*
1118	2468	A	9154	471	2	GILTKLSAFLTIPRLQPHLIAALSPSS AAGQVVVEVTSHLYLCITSDAAGLRLLPPAES ERGEGGHCPAEAPLPPRPQYCLAKHPLLRKLP EEKIKLDPYLTQHTKINSKQIKYLS/VRAKTTQ LVEGNIGVNLQNTELKQH*INGFLDTTPEAQE TKEKTNKLNFIKKVKRQLAEWEKIFQIA
1119	2469	A	9155	-	3187	ACPRLARRRRVVSLRRRRGWLRARWSRGQ NNMAARRITQETFDAVLQEKAKRYHMDASG EAVSETLQFKAQDLLRAVPRSRAEMYDDVHS DGRYSLSGSVAHSRDAGRESLRSDVFSGPSFR SSNPSISDDSYFRKECGRDLEFSHSNSRDQVIG HRKLGHFRSQDWKFALRGSWEQDFGHPVSQ ESSWSQEYSFGPSAVLGDFGSSRLIEKECLEK ESRDYDVDHPGEADSV/LRGGSQVQARGRAL NIVDQEGSLLGKGETQGLLTAKGGVGKLVTL RNVSTKKIPTVNRITPKTQGTNQIQKNTPSPD VTLGTNPGTEDIQFPIQKIPLGLDLKNLRLPRR KMSFDIIDKSDVFSRFGIEIIKWAGFHTIKDDIK FSQLFQTLFELETETCAKMLASFKCSLKPEHR DFCFFTIKFLKHSALKTPRVDNEFLNMLLDKG AVKTKNCFFEIIKPFDKYIMRLQDRLLKSVTP LLMACNAYELSVKMKTLSNPLDLALALETTN SLCRKSLALLGQTFSLASSFRQEKIL*AVGLQ DIAPSPAAFPNFEDSTLFGREYIDHLKAWLVS SGCPLQVKKAEPEPMREEBKMIPPTKPEIQAK APSSLSDAVPQRADHRVVGTIDQLVKRVIEGS LSPKERTLLKEDPAYWFLSDENSLEYKYYKL KLAEMQRMSENLRGADQKPTSADCAVRAML YSRAVRNLKKKLLP\WQRRGLLRAQG\LRG\ WKARRA\TTGTQTLLFLRAPGLKHHGRQAPG LSQAKPSLPDRNDAAKDCPPDPVGPSPQDPSL EASGPSPKPAGVDISEAPQTSSPCPSADIDMKT METAEKLARFVAQVGPEIEQFSIENSTDNPDL WFLHDQNSSAFKFYRKKVFELCPSICFTSSPH NLHTGGGDTTGSQESPVDLMEGEAEFEDEPP PREAELESPEVMPEEDDEDDEDGGEAPAPPG GAGKSEGSTPADGLPGEAAEDDLAGAPALSQ ASSGTCFPRKRISSKSLKVGMIPAPKRVCLIQE PKGECPPVGTVASSTVLGWWAVRVRRDRWR HFNPKEFCAPLQNVSRHSCFPVV
1120 1121	2470 2471	A	9163 9166	124 272	207 523	PPRACRPCPRACPCPPT*KCSQPVSWPC PMSSLQGCFYTFKCIIFKGIFLLLISNLIAF**EK
						V/CSHITDSLKFIGKGWVGMVTHACNPGTLG G*GGWIA*VREFETSLGNM
1122	2472	С	9170	442	236	MNRRRFLRPADCHSGMRGTENGACSEGESQI HCGAGGEGVQLVHVVNQPENGCLQFDSTHIT FSKRQN*
1123	2473	Α	9171	10	423	MVDRSPLLTSVIIFYLAIGAAIFEVLEEPHWKE AKKNYYTQKLHLLKEFPCLGQEGLDKILEVV SDAAGQGVAITGNQTFNNWNWPNAMIFAAT VITTIGYGNVASKTPGGRLFCGFYGLFGVPFC LTWINALGKFFG

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			•	amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			ĺ	peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence	ļ	nucleotide insertion
1124	2474	A	9173	3	374	GPSPSLLVLLPQEPGGTGTPVRAGAGAGMWL
	ĺ					WEDQGGLLGPFSFLMLMLLLETRNPVNACLL
		İ	ŀ			TGSLFVLLGVFSFEPVPSCRALQELKPRDRISA
	<u> </u>					IAHRGGRHDPPENTLGAIR/QGS**WSNRR
1125	2475	A	9179	704	188	ESSSGLLFQCFQGIHVQKLTLQARPTLFSWWL
			1			CSKPPKETGELENAESGGDGGRRGGKQDNV
		1	ŀ			AWWRRM\QKG\DFPWDDEDFPQSGPFGGQA
						LPMGFFYLYFRDPGREITWKHFVQYYLARGL
	ĺ	ĺ	Í			VDRLEVVNKQSVRVIPAPGTSSEVRGEFKAE YCRHKFISCKNVVFYFFQ
1126	2476	A	9183	153	233	MEYMAESTDRSPGHILCCECGVPISPN
1127	2477	A	9185	1	321	LTGQLGSILLRVFSKSRAGLGARKLKAYRTM
		1	7105	ļ -	321	EYMAESTDRSPGHILCCECGVPISPNPAQY\CV
				1		ACLRSSFHIYHCIPKLFIHPFSKTSSSAFITPSHY
	•					LTFFSTIS
1128	2478	A	9186	183	847	VLKFLLLQTMDEQSQGMQGPPVPQFQPQKAL
		1				RPDMGYNTLANFRIEKKIGRGQ\FSEVYRAAC
	ļ					L\LDGVPVALKKVQIFDLMDAKARADCIKEID
		l			·	LLKQLNHPNVIKYYASFIEDNELNIVLELADA
		i	-			GDLSRMIKHFKKQKRLIPERTVWKYFVQLCS
						ALEHMHSRRVMHRDIKPANVFITATGVVKLG DLGLGRFFSSKTTAAHSLVGTPYYMSPERIHD
	Ì	1				NG
1129	2479	A	9190	1	370	GTSWKIPSAAVSESSPNGAAYASGLPCGVRG
	1	_		_		PPWAGLALLPSPTLMALLRRPTVSSDLDNIDT
)				RATT\KIRVVATITRARIEDMRHSATALTRPD
						ATTAQIPKLPVTTVCNRRANPGIPPSVL
1130	2480	A	9194	131	487	AYLKRLPVPESITGFARLTVSEWLRLLPFLGV
			1			LALLGYLAVRPFLPKKKQQKDSLINLKIQKEN
						PKVVNEINIEDLCLTKAAYCRCWRSKTFPAC
1131	2481	A	9201	184	605	DGSHNKHNELTGDNVGPLILKKKE
1131	2401	^	9201	104	003	KELVDEKSERGRAMDPVSQLASAGTFRVLKE PLAFLRALELLFAIFAFATCGGYSGGLRLSVD
	ĺ				Ì	CVNKTESNLSIDIAFAYPFRLHQVTFEG\PTCE
						GKERHKLALIGDSSSSAEFFGTVAGFAFLYSL
		1				AATGVYIFFQNKY
1132	2482	A	9206	1	852	GGGRAGAGSRDMGSTDSKLNFRKAVIQLTTK
						TQPVEATDDAFWDQFWADTATSVQDVFALV
						PAAEIRAVREESPSNLATLCYKAVEKLVQGA
						ESGCHSEKEKQIVLNCSRLLTRVLPYIFEDPD
		[WRGFFWSTVPGAGRGGQGEEDDEHARPLAE SLLLAIADLLFCPDFTVOSHRRSTVDSAEDVH
		1				SLUSCEYIWEAGVGFAHSPQPNYIHDMNRME
	[LLKLLLTCFSEAMYLPPAPESWQH/RTHWFSS
		1				FVSSENRHALPLFTSLLNTVCAYDPVEYGIPY
			1			NHLY ·
1133	2483	A	9208	1165	1463	GPRARVQGFSGADIVKFMALGSMYLVLTLIV
]	AKVLRGAEPCCGPLKNRVLRPCPLP/VPLPPP
		l				HPQPSRGNPVGCLPTYKVVYKLLSWPLHSNS
1124	2494	 	0210	66	1506	NVYFIV
1134	2484	Α	9210	66	1586	MAGAGPKRALSAPVAEEKEEAREKIMAAK
						RADGAAPAGEGEGVTLQGNITLLKGVAVIVV AIMGSGIFVTPTGVLKEAGSPGLALVVWAAC
		1				GVFSIVGALCYAELGTTISKSGGDYAYMLDV
		1				YGSLPAFLKLWIELLIIRPSSQYIVALVFATYL
		ł				LKPLFPTCPVPEEAAKLVACLCVLLLTAVNC
	}					YSVKAATRVQDAFAAAKLLALALIILLGFVQI
	l		1			13 MAATIK QDALAAAKEEAEAEIIEEGF VOI
						GKGDVSNLDPNFSFEGTKLDVGNIVLALYSG LFAYGGWNYLNFVTEEMINPYRNLPLAIIISLP

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
				- -		IVTLVYVLTNLAYFTTLSTEQMLSSEAVAVDF
		İ			-	GNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSS RLFFVGSREGHLPSILSMIHPQLLTPVPSLVFT
]			CVMTLFYAFSKDIFSVINFFSFFNWLCVALAII
•			}			GMIWLRHRKPELERPIKVNLALPVFFILACLF LIAVSFWKTTPWSVASDFTIILSGLPVYFFGV
						WWKNKPKWAPPGHLSPRPSCVRSSCMVVPQ
1135	2485	A	9216	40	410	RDRLPPAYFCRPVVCVVTALDVG\SPESQEM
						DLVAFEDVAVNFTQEEWSLLDPSQKNLYREV MQETLRNLASIGEKWKDQNIEDQYKNPRNNL
]]			RSLLGERVDENTEENHCGETSSQIPDDTLNK
1136	2486	A	9223	3	983	RRRRSRYRRCSRFPRPGPLAVSMPHAFKPG
						DLVFAKMKGYPHWPARIDDIADGAVKPPPN KYPIFFGTHETAFLGPKDLFPYDKCKDKYGK
			ł)	PNKRKGFNEGLWEIQNNPHASYSAPPPVSSSD
İ	1	ļ				SEAPEANPADGSDADEDDEG\RGVMAVTAVT ATAASDRMESDSDSDKSSDNSGLKRKTPALK
į						MSVSKRARKASSDLDQASVSPSEEENSESSSE
						SEKTSDQDFTPEKKAAVRAPRRGPLGGRKKK
						APSASDSDSKADSDGAKPEPVAMARSASSSSS SSSSSDSDVSVKKPPRGRKPAEKPLPKPRGRK
']	PKPERPPSSSSSD
1137	2487	A	9229	21	239	LFPRLECRDPVTVNCTLNLPGSKNAPTTASQV
ļ			}			GSTWNYRGGLPHPTNFFVKTGFRCSQAGLKL RGSREPPAWA
1138	2488	A	9231	1664	2	TRSVGVNTCEVGVVTEPECLGPCEPGTSVNL
ľ						EGIVWHETEEGVLVVNVTWRNKTYVGTLLD CTKHDWAPPRFCESPTSDLEMRGGRGRGKR
						ARSAAAAPGSEASFTESRGLQNKNRGGANGK
	1				1	GRRGSLNASGRRTPPNCAAEDIKASPSSTNKR
						KNKPPMELDLNSSSEDNKPGKRVRTNSRSTP TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK
						YKHINGLRYHQAHAHLDPENKLEFEPDSEDK
						ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA
			[SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA
						DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK
						ANNCKTDKN\PSKLKSARPIAPAPAPTPPQLIA IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL
					[KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR
				·		KLKDKEGKETGSPKMDAKLGKLEDSKGASK
						DLPGHFLKDHLNKNEGLANGLSESQESRMAS IKAEADKVYTFTDNAPSPSIGS
1139	2489	A	9234	207	443	TRRGQPWRRRAAAAGILPGREAAACLPSC/AS
						VTAAVSGLLVGYELGIISGALLQIKTLLALSC
1140	2490	A	9238	248	328	HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV
1141	2491	A	9242	2	535	FVEAAVKMLGSLVLRRKALAPRLLLRLLRSP
						TLRGHGGASGRNVTTGSLGEPQWLRVATGG RPGTSPALFSGRGAATGGRQGGRFDTKCLAA
						ATWGRLPGPEETLPGQDSWNGVPSRAGLGM
			1			WPWAAALVVHCYSKSPSNKDAALLEAARAQ
1142	2492	A	9245	157	466	WMQEVSRNRCALLHSAAVQEYGYGN HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD
1172	2432	^	7243	157	100	FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC
			}			CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF
1143	2493	Ā	9247	264	115	GICKEYSRQ GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG
						ARDSTSIIRMGPEIPPP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1144	2494	A	9260	1	401	KKVPGRLSEMSFSLNFTLPANTTSSPVT\DCGP SLGLAAGIPLLVATALLVALLFTLIHRRRSSIE AMEESDRPCEISEIDDNPKISENPRRSPTHEKN TMGAQEAHIYVKTVAGSEEPVHDRYRPTIEM ERRR
1145	2495	A .	9264	175	411	METIWIYQFRLIEIGDSTVGKSCLLHRFTQGRF PGLRSPACDPTVGVDFFSRLLEIEPGKRIKLLL WDTAGQERFISIT
1146	2496	Α	9277	592	814	MFTYLEGREGIKSQPKMEPHSVTVRLECSGMI SAHCSLNLPGTSDSPASASR/VAGTTGMRHHA WLIFAFLVETGF
1147	2497	Α	9279	1255	2	FRRGRRGEEEKEEEEEEEGWVNGMENSHPP HIHHQQPPPQPGPSGERRNHHWRSYKLMIDP ALKKGHHKLYRYDGQHFSLAMSSNRPVEIVE DPRVVGIWTKNKE\LELSVPKFKIDEFYVDQV PPKQVTFAKLNDNIRENFLRDMCKKYGEVEE VEILYNPKTKKHLGIAKVVFATVRGAKDAVQ HLHSTSVMGNIIHVELDTKGETRMRFYELLV TGRYTPQTLPVGELDAVSPIVNETLQLSDALK RLKDGGLSAGCGSGSSSVTPNSGGTPFSQDTA YSSCRLDTPNSYG/QGTPLTPRLGTPFSQDSSY SSRQPTPSYLFSQDPAVTFKARRHESKFTDAY NRRHEHHYVHNSPAVTAVAGATAAFRGSSD LPFGTVGGTGGSSGPPFKAQPQDSATFAHTPP PAQATPAPGFR
1148	2498	A	9302	1026	6	IASIQNADTMPGVGLLVSHFSTLVSRQRCPNY ADPQNLTDVSIFLLLEVSGDPELQPVLAGLFL SMCLVTVLGNLLIILAISPDSHLHTPMYFFFSN LSLPDV\GFTSTTVPK\MIVDI\QSRSRVISYAG CLTQKSLFAIFGGTEE\NMLLSVMAYDRFVAI CHPLYHSAIMNPCFCAFLVLLSFFFLSLLDSQL HSWIVLQFTIIKNVEISNFVCDPSQLLKFACSD SIINSIFIYFHKDPERQLVLAGLFLSMCLVTVL GNLIIILDVSPDSHLPTPMYFFLSNLSLPDIGFT STTVPKMIVDIQSHGRVIFYAGCLTQMSLFAIF GGMEERHAPECDGL
1149	2499	A	9303	1	699	MASQEKDIFIGWGTIHLFRKPQRSFFGKLLRE FRLVAADRSMGRYMLFGVINLICTGFLLMWC SSTNSIALTSYTYLTIFDLFSLMTCLISYWVTL RKPSPVYSFGFERLEVLAVFASTVLAQLGALF ILKESAERFLEQPEIHTGRLLVGTFVALCFNLF TMLSIRNKPFAYVSEAASTSWLQEHVADLSR SLCGIIPGLSSIFLPRMNPFVLIDLAGAFALCIT YMLIEI
1150	2500	Α	9308	797	693	DRSTSVTRAGVQWCSLGSLQPRTPGLLRSSCL SLP
1151	2501	A	9309	205	406	VAIKELPVLWKWSKPTR\TAKEPPQTQQRAG SKTAAPPCQWSRMASEGPNIPCPGARHSDKQ FLICTI
1152	2502	A	9314	913	504	KPSPLITPPAVVLPPSAVLNLVNTFSSFPQVEV QGPLCGPRKGRLAVTIPFFGLS/LPKYMDHRR PPPHR\EIFFVFLAETGFHRASQAGPDLPTS/S/I PPTSA/FPKCWEYRSEPQCLPGCLSFSGILLDL GTNVSLRAA
1153	2503	A	9315	392	1	HPHRPRPGFRSPARSSRPCPVLTSLLPPFPSPSP PADDLVKAGRDRKDPQVR/ERRLRPNPGRLG GPR\PRPARARS/CHQPRLTRVCPRSPPPEARA PAPAAPARGRGAPKRNRPRTDTRAPRGSSAR PGNS

COTO ID	LCEOTO	Met	1 000	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	SEQ ID NO: of	hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ſ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
nence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
201.00		1	***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide		/=possible nucleotide deletion, \=possible
		1		sequence		nucleotide insertion
1154	2504	A	9321	331	433	MPCI/QAQYGTPAPSPGPRDHSASDPLTPEFIK
		ĺ				PT
1155	2505	A	9324	180	275	MEEPQSDPSVEPPLSQETFSDLWKLLSENNVL
1156	2506	Α	9326	383	619	MISPSRTEGDPLPLPP/EGEGQEVRGFGGGPAK
		j			<u> </u>	EAAQRHCRASVSILRMRRPGQGSSRPARVPL
	İ	l				RGPDSHRLREPPPSPP
1157	2507	A	9327	152	292	YERRGRSQGGSHPAGAQPGGRAIGAGWQS
		Ľ				KEPLWEGLQRSGSPLPG
1158	2508	A	9328	1	430	QELKQGPNPLAPSPSAPSTSAGLGDCNHRVD
		İ				LSKTFSVSSALAMLQERRCLYVVLTDSRCFL
		l				VCMCFLTFIQALMVSGYLSSVITTIERRYSLKS
0.0		j				SESGLLVSCFDIGNLVVVVFVSYFRGRRRRP/
		<u> </u>				RVAAVGGLLDLEGGEMI
1159	2509	Α	9334	108	383	KGNQVNGNGNQLKRKHESMCPVSLTQNTVR
		1				LMEAGLPQKQAERADELFEAGLVIYVKLDER
11.50	0510	-	0000	<u> </u>		VLNAL\YSSVGLQWFKESDLSHLRLLEISFR
1160	2510	Α	9338	2	430	FVGRPRGLSDRLEDLFLAGFRVGERLRTAAM
						KRYVRILLLGEGAEHVADPVPGGRGVPRGEA
						DHTDQELREEIHKANVERVVHDVSQEATIEKI RTKWIPLV/RWGDHA/EGPVGIKSYLPSGRSM
		1	{			
1161	2511	A	9341	1	390	EAELPIMSQLTEIETCVEC NSRVDDFVAPGLSEAGKLLGLEFPERQRLAA
1 1101	2311	Α.	9341	1	390	AVG/CSPMSGVISMSAPFFLGKIIDAIYTNPTV
						DYSDNLTRLCL GLSGVFLCGAAANAIRVYLM
		l	•			QTSRQRVVKRLRTSLFSSILGQEVAFSDKAGT
j		j]		1	GELI
1162	2512	A	9343	84	837	QGRFRAFCWQRDFLQPPGMRLSALLALASKV
			1 2	* '	***	TLPPHYRYGMSPPGSVADKRKNPPWIRRRPV
		1	9			VVEPISDEDWYLFCGDTVEILEGKDAGKQGK
		Ì			ì	VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR
						GTMIPSEAPLLHRQVKLVDPMDRKPTEIEWR
		İ				FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET
		l				WIDGPKDTSVEDALERTYVPCLKTLQEEVME
		1	İ			AMGIKETR\NTRRSIGIEPGAEQLLPNFCPSLE
						G
1163	2513	A	9346	967	616	DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS
		ł	į	ļ		LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ
		ł				AGLKLLISSDPPTSA/FPKCWDYRRD\SSAPAT
						FSSYQRNNPDLILNDTIMPNIK
1164	2514	A	9347	3	1099	SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI
		1 .			ĺ	HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL
		1				FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV
		1		[VAGASVGAGVWARNPRYRTEGEACVEFKA
		ļ	ļ	j]	MLIAVGIHLLLMFEVLVCDRVERGTHFWLL
						VFMPLFFVSPVSVAACVWGFRHDRSLELEILC
		1				SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM
		1				SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH
l	1	l	1	ł		VTMAISWITTVVPLLTFEVLLVHRLDGHNTFS
	1			ļ	1	YVSIFVPLWLSLLTLMATTFRRKGGNHWWF
	1	1	1]	AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA
1165	2515	-	9362	547	001	LPLQNKDRGSWPASRGSPRLL
1100	2313	A	9302	347	991	DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP VPEGVRLADGPGHCKGRVEVKHQNQWYTV
		-		1		VPEGVKLADGPGHCKGKVEVKHQNQWYTV CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC
		1				TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP
]	1	}	J]	LGEDTLFHVEYTSVHGRERLSAKD
1166	2516	A	9363	201	387	PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS
1100	2310	^	2005	201	36/	GAISAHLAHCNLCLPGSSDSPASAFQVAS
1167	2517	A	9368	707	1087	AVLTPCLSPCSPSRIPRP\SRPYPGRRSLSHTPP
110/	1231/	<u> </u>	1 2300	101	100/	WATER CONTRACTOR LACKESTOLIAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PRPLILYAPAP\RPAGTAFIPHSHPPPPDLLRPT
			•	:		ATPA/TPCPSLPPPPRPLHPTQPSTALLPDPPPW PLPFPPPSS/RPPRPDCSTSYSPTFPPPT
1168	2518	A	9375	511	15	MMLSEETSAVRPQKQTRFNGAKLVWMLKGS PITVTSAVIIVLMLLMM/IFSPWLATHDPNAID LTARLLPPSAAHWFGTDEVGRDLFSRVLVGS QQSILAGLVVVATTGMIGSPLECLFGELGGRA DAIFMRVMDIMRS/IPSLVLTMEKTAALGPSL FNAMQASSEH
1169	2519	A	9377	42	410	GNGRVAPRDPGAVASAEPGLTTHDSGVNPN NSARRMEAMASGSNWLSGVNVVLVMAYWS LVFVLLFIFAKRQIMRFAMKSLRGPHGPVGH NAPKDLKEEIDILLSRVHNIKYEP\HLLADDDA
1170	2520	A	9378	302	1303	GVSGFSASVLRQRRMEDELEPSLRPRTQIQGR ILLLTICAAGIGGTFQFGYNLSIINAPTLHIQEF TNETWQARTGEPLPDHLVLLMWSLIVSLYPL GGLFGALLAGPLAITLGRKKSLL\VNNIFVVS AAILFGFSRKAGSFEMIMLGRLASWGVNAGV SMNIQP\MLPGGESAPKELRGAVAMSSAIFTA LGIVMGQVVGLSTTAATGLRGL\AGELEELEE ERAACQGCRARRPWELFQHRALRRQVTSLV VLGSAMELCGNDSVYAYASSVFRKAGVPEA KIQYAIIGTGSCELLTAVVSVSLEGALPPPAL WGGTPRSFALNQFTLQKKKK
1171	2521	A	9381	2	412	RGPASAQEDERARTAPLERVRARGRMTTSSA LFPSLLPCSWSTSNKYLAEFRAGKMSLKGTTE TPDKRKGLAY/IQQTDDSLIHFCWKDRTSGNV EDDLIIFPDDCEFKRLPQCPNGRVYVLKFKAG SKRLFFWMQEP
1172	2522	A	9384	20	355	GWNGRSTEASPAAEAPHVPHKETKAAMGTQ CTHGGKVRPDPHDMLTTVVHKIKLFVLCHSL LQLCAIMISDYLKSSIYTVEKRLGLFRPTSGLL ASFNEVGNTALIVLESY
1173	2523	A	9393	430	87	LCQCIVPGQQKETFSLNPSSATVRFYL*LSLQ QRKEDQ*IIL*YHLNKDCLHIFMSAITLYMKI* KIFVLFDFNIMFETPFYII*FIFLFSQNLKRIRQV IRPPISFSKINNGP
1174	2524	Α	9397	77	374	ERLEIGRLGGERGSGPASCLRVIDVSGMWDQ RLVKLALLQLLRAFYGIKVKGVRVHRDCGTF ESSSTLIRVS*FGVPCNALAHFGVTHF*YILDF LGML
1175	2525	A	9399	66	397	HESSRADRDKMDTRGSTYTDADPVNKSGGT AKMNKWSKGKVRDKLNNLVLFDTATYDKL CKEVPNYKLITLAVVSERLKIPGSLARAALHE LLSRGLI*LVIQHIAQVIY
1176	2526	A	9408	2	299	LDLTHVLSLSISLTVTILGTTFGMVIPLLDVVY GERGYAQNGDF*DAQLDDYSFSCYSHAQVN GAPNSLTRAYDDP*VKISGLECQKVGALVEV KCLNL
1177	2527	A	9416	2	402	CNFLRSSRIRVHSTPAASTMPPKVDPNEIKVV YLRCTGGEVRATSALAPKIGPLGLSSIKVGVD FV*ATGDWNVLIISVILTIRILLSHIFVVPPFFCF DHLIAFWDLQSLIFLHVIFSLFITLLLFCFFSIF
1178	2528	A	9419	142	426	TPLFDLWPRVVLSWLETVLTSLRTRRAASGPP ACRIMPTTVDDVLEHGGEVHFLQKQMLYLL ALI*DTFAPIYVGIVFLGFTPDHRCRSPGVAEL
1179	2529	A	9420	1450	1655	LSSAGTKMNLN*KNYWPGASAHACNPSTLG GQSRCITRSGDRDHPG*HGETPSVLKIQKISRA WWRAP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	j	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	dence	Į	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence	1	İ	714			T=Threonine, V=Valine, W=Tryptophan,
1	1	1	i	amino acid	of peptide	1=1 freonine, v=valine, w=1 ypiopiiai,
	}	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
J	}	J	1	peptide	J	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1180	2530	A	9422	176	375	HRPQTTRPDWKPRT*PQGK*GRLSSEISPASPP
1		١			***	SRFSRSTKPVPPKADPPARQKLTGVLHAPLLK
	1	i	1			L
		↓	0426		274	PIAASLRMYNLQPYTEENLICTAFATMVETVP
1181	2531	Α	9436	2	1 2/4	
ļ	Į	ļ.	}	ļ	ļ	IARTILDRLTGIPHGYCFVE*ADWATADKCVH
İ			l .		1	IYNGKPLPGATPLLSLQLHQLAHLGS
1182	2532	A	9442	3	240	VDKCSSKSIVLSEYCPHCMCSLSTDPKPFGQL
						SMILK*MGAGDEKISAMGKARVDHRELYLGL
	ł	}				LYPTEDYKLTFRARH
1102	0522	 	9444	384	3	LKDFOPWALHDWPLFCCCTFLLFLVLECFTR
1183	2533	A	7444	204	, ,	
1					1	KGCSGWAPWLSLQCQHFGRPRWADHLRSGV
	}	1	1	!		RDQPGQYSKTTFLPKIQKLAGHSGAHL*S*LL
	ŀ	ļ		· ·		ERMRWKNRLNPGGRSCSEPRWHHCTPGWAT
		l			İ	ERG
1184	2534	A	9462	391	655	LSGFKSLMPKIPLQYIYVRVRTTWSFCLPLDG
1107	2234] ^	7702	""	1 333	RKLMLS*YSK*LT*KYNILPEYSRMTLPPGMV
				1		IHTCNPSTLGGRAGWIV*AQEFET
<u> </u>		<u> </u>				
1185	2535	A	9467	215	566	RCPMWQGQASRMDPAKAKDREASTCCSLA
1			ì	}	ļ	WWWGWECWVRALKLSSGPAGPLACWVAK
1						KKSLSLSGPVYPSEKGAGLYVF*DRVSLCHPG
	1	İ				WSAVVQFWLTAASNSCFSLLSSWDYRCA
1186	2536	A	9468	275	452	HIPQLHTKTHYVPTRMVNKI*QIDNSKPWQR
1100	2330	^	7400	2/3	432	GG*TGILTHCW*ESKLVQPLWKIVWHYQ
	L		L			OG. IGILIACM. ESKLYQILWKIY WITTQ
1187	2537	A	9469	388	3	EVAPGPSQILPRRVTLGGDRPQFSLPGPRLPQ
	i	1				SSRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP
1		l		1		NPASPHPEAPQEPWDSASGSVGSFSLGRGAK
					}	ASS*VPGKGRGPRQGSELLAETILELFLALAN
J	j	j	J	i	1	S
1188	2538	A	9471	124	397	TMDKKNRHGNSLDMASEIHMTGPMCLIENTT
1100	2,338	^	74/1	127	377	GRLMANPEALKILSAITQPMVEEAIAGLYRAC
						*FYLTNNLAGMKKGLCLGSTEQAHTIGI
1189	2539	Α	9480	584	769	GHVQSQHFGRPRRADHLRSGDRDHPG*HDET
		1		1		PSLLKIQKISWAWWRAPVVPATWEAEAEEW
1	1	1	ł	1		R
1190	2540	A	9483	463	86	VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL
11150	2540	1 ''	1 7 105	1.03	••	PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR
I	j	İ	1	i		
1	1	1		1		GASSCRRRCNPVLAARKAGSPRSHSTRENC
	<u> </u>					RRSRCPDTAHRRRRRGRRRNPSCVRSPRWR
1191	2541	Α	9489	1	411	LADALCLSAAATGAVRPGARAQPSTRRRLSP
1	1	1 .		1		SVRVCCRAAAASNLLYSSCLQRHSERASEEG
1	1	1		1		ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI
1	1	1	1	!		MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ
1	İ		1	}		KEEELTAVNVK
1100	0545	 	0405	200	161	
1192	2542	Α	9497	389	161	VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI*
		1		1		CEEDERKMAREFLAEFMSTYVMMNIHMIVE
		1	.L	<u> </u>	<u> </u>	KDTYSDHEENTS
1193	2543	A	9509	186	1	IAKSQ*KRWQRSGAMETLKHGWWECKLVQF
1	1	1	1			FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM
1194	2544	A	9512	58	433	PLORSKCLTLRCLRAKPWAWSQSPRACSSAL
1174	2344	1^	7312	30	7.55	
1	1	1	1	ļ		LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA
1		1	I			SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI
	1	1	L			RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ
1195	2545	A	9515	595	1223	GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP
		1 -	1			AGGGASLPVAAGSCAAAPHTEPGAPOHLLDC
1		1	ŀ			PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP
1		1	1	1		LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL*
}		1	1	ļ		
						PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP
		1	1			APGSGPCGATARPSRGGRAGGSRARRPIPPGP
		1	1	ļ		GTRRTPSGCQNPAASGG
		•		`		

CEO ID	SEO TO	Mark	SEO.	Dradiate 1	Predicted end	I American de la Companya de la Comp
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in NO:	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-	1	USSN	location		F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence		09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	-355	ļ	914	ng to first	acid residue	Q=Glutamine, N=Asparagine, P=Proline,
		[amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	,	l		peptide	1	/=possible nucleotide deletion, \=possible
l		ļ	1	sequence	1	nucleotide insertion
1196	2546	A	9518	229	468	RSPTATPAPHAMGPGAPFARGGRPLPLLGAM
					1	AERVAPGWDLHTPYLPRTNSRRTPHL**EPHA
						GYIGALFPMSGGWPGGQ
1197	2547	A	9521	289	448	IAWLSGLFFPSNQANLCFLCYKLTADSRYRG
	[1]			HAMRHLTGNTSMAIRFL*ADSRFQVQRARYE
<u> </u>	0515	<u> </u>			<u> </u>	APNWKYKYGY*IPVDMLC
1198	2548	Α	9524	204	1	KNKKTTKCLSIVTLNISGPNQ*NKRHRVAEWI
	[1	VKQEPNICHL*ETHFPFRDTYRLKEREQKKRK
1100	2540		0515	1505	10.45	SSYS
1199	2549	Α	9546	1785	1943	GGRFKESKLTNAGWQRNSFFIGPPKSIPWAA
1200	2550	_	0540	104	 	V*QRGDGKNPGVTHLNRPVGTX
1200	2550	Α	9548	186	1	VNAEKEF*KIQHYFMTKSQNKLHIEHTYLKPI
1201	2551	<u> </u>	0540	501	 	KAIYDKWTSDIMLNLQKL*AFFLRVIVRQI
1201	4551	Α	9549	591	2	SSVVEFPRGPRSSLPPLDSTFPCGSSPNWTGGC
1	[Į.				GSCPSGE*LVSPGSEQRKKYSNSNVIMHETSQ
	[1	1			YHVQHLATFIMDKSEAITSVDDAIRKLVQLSS
			j 1		•	KEKIWTQEMLLQVNDQSLRLLDIESQEELEDF
	1		 	ļ	l	PLPTVQRSQTVLNQLRYPSVLLLVCQDSEQSK
		1				PDVHFFHCDEVEAELVHEYMESALTDCRLGK AMRP
1202	2552	A	9552	428	1	
.202	عدده	\ \frac{1}{2}	2000	720	'	KYGNEGHWSRQCPNPGKPIRPCPLCRGPHWK
		1]	ļ	l	LDCERPPQGPLPSLPELAKTSYSDLTGLATED *WGPGMDAPATTIASSYTPYTIAYAGPPYEE
		1				*WGPGMDAPATTIASSKTRVTLMVAGRPVFF
		'				LI*YRATYSALPNFSGPTQSSQVSVVGIDGQV SKPRATPPLFCSLHTF
1203	2553	A	9568	517	738	RRKFERKQKQ*RYREGKQYRQRDKMKEWG
			200			EKEKRREKGEREERKMRHRERKGESGQRD
			.			TMENWRVERLTEKER
1204	2554	Α	9573	83	415	EDKRLRLVDGDSRCAGRV*IYHDGFWGTICD
]		-			DGWDLSDAHVVCQKLGCGVAFNATVSAHFG
				1		EGSGPIWLDDLNCTGTESHLWQCPSRGWGQ
						HDCRHKEDAGVICSEFTALR
1205	2555	. A	9577	64	424	ARGSCPTRPRTANGRMGETKDAPQMLVTFK
		!		1		DVAVTFFREEWRQLVLVHRTLYR*GMLETC
			ļ l		, "	GLLDTLRHNVPQPDVVHLLYHGTQLLIVKRE
	·	<u> </u>				VSHSPCAGDMRELFTREATLTPHPYNNGA
1206	2556	Α	9584	38	476	TLGAVLFSEVSKESSTSHSGGQLGRQNRHPKL
	į l				1	SNFITPSSPRLKP*TASSQRNLGQILNMFLTAV
	}	 	, . I		1	NPQPLSTPSWQIETKYSTKVLTGNWMEERRK
	ļ l		į į		1	GLPYKHLITHHQEPPHRYLISTYDDHYNRHG
10	المستا	أحصيا	L			YNPGLPPLRTWNGQKLLWL
1207	2557	Α	9586	2	412	LRSSPAALLRALCITTVTGTALALRSRVATTN
			į į	1	1	PDGCRNVLRPKYYRLCDKAESWGIALETVPT
İ	1 1	ا ا	l l			GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP
	ļ l	l	ļ l	'		THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL
1000	0					FGILFSICFS
1208	2558	Α	9597	122	3	IKNYWPGMVAHACNPSPLGGRGRWIA*AQK
1000			لــــا			FADAWADAW
1209	2559	Α	9611	148	558	KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK
I	ļ l	Ţ į	į l	İ	' I	GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ
		ļ l			' I	RIRDHDLLDKRKTVTALKAGEDRAILLGLAM
Ì	ļ .	1 1	1]	' J	j	MVCSIMM*FLLGITLLRSYMQSVWTRESQCT
1210	25/0	اــــــا	06:2	204		LLNASITETFNC
1210	2560	Α	9618	384	2	SLHDMLMLAEQQQKQKWAVNTQNTAWSNA
		ļ	1		' 	DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI
			1. 1		·	KVQVKNNDLGLQATINNEANWIAHQDDFNW
l		!	ı İ			LLAELNTCQRQETADS***WSPKNSHVGKDS
1211	2561	<u> </u>	9620	316	610	GELSAK
	2701	Α	7020	316	610	QKHPGGGQLGRSPQEDSRFHNKASSGVSRVR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LGRAWWLTPVIPTLWEAKAGGSPE*D*AGRG
1212	2562	A	9623	297	344	GSRL*SQHFGRPRRVDHLRSAVQDQPGQHGE TPSLLKIQKIN*VWGRRL*SSYSEAEAGESL QFPVDGDYQKIEKITQLFQAQNLSLCLAMTR TREL*KGGGKGRHE*AVVPFLKKGGYGVKAP AILNTSNCT*CF*ETKMLSDDPKACVFEVSSA
1213	2563	Ā	9624	2	356	DL*NTSFGVIR AELSLASTACGRNTSGDSLPDYDRAPISSPLA TSGTILSAISCLWDLPTPVLRVGLSCQPSMSSQ IPRMYSTDVEAAVNSLEDLYLQAYYAYLCVG LYFHRDDMALEGVSRFL*ELAE
1214	2564	A	9634	776	912	SLSRWVRAKL*VPYNQENCLNPRGGGCSEPR SHYCTPAWATEKDS
1215	2565	A	9636	220	426	KPGNFAVSSEY*DITSGQLKTAVRG*IEMTST EENFGEKLHDIGFGNGFLDKT*KAQATKAKI DK
1216	2566	A	9637	391	76	CFLEDGCTQAS*AEEAAVSPSMAEEEQGSTSC RERRSIRFKMKNHSPDDTIKENVTISNIRTRKI NHLPETERNLLEHGLMYIRLNAAFCSLVAHS LFGFILKAT
1217	2567	A	9655	2008	2432	LHCKMGALETQTHPCSQNMLRSLQKCCCKV EEHHLQPVQVLQTLLHSATAGTGCRRPARPP PAPPTPTPWRSRQSGKQSERAS*LKGRGRYGL GALGGRGGRALGGSRWPP°LPGETLFSGCKH RRRRGSDAAPGEEAGT
1218	2568	A	9658	3	405	HASARALLSPNLSPNNKMAISGGPVLGFFIIA VLMSAQEPWAIKEEHVIIQAEFYLNPDQSGEF MLDFEGEDTFHGDMAKKETVWRLE*LARLD NFEAQRALANIAADQAALEIMDMGSDYTLIP NVPPKVTVL
1219	2569	A	9662	3	284	PDWTEKRKMQDTGSILPLHWFGFGYAALVA YGGIIGYVKAGSVPSLAAGLLFGSLSGLGAYQ LSQDPRNVWVFLATSGTLAGIMGMRFYHSG KL
1220	2570	A	9669	200	699	LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL K
1221	2571	A	9676	164	562	KERDSSTFSAAMTTMQGMEQAMPGAGPGVP QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV VQILTALMSLSMGITMMCMASNTYGSNPISV YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG SLGMNITSS
1222	2572	A	9688	43	412	VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF PTDENIKRKWVLAMKRLDVNAAGIWEPKKG DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS PYHLQGKREKLHCRKNFTLKTVPATNYNH
1223	2573	A	9696	308	564	RTSMGILYSEPICQAAYQNDFGQVWRWVKE DSSYANVQDGFNGDTPLICACRRGHVRIVSFL LKKECLCQPQKPERENLLALCCE
1224	2574	A	9700	3	632	DAWASGELGSLFDHHVQRAVCDTRAKYRE GRRPRAVKVYTINLESQYLLIQGVPAVGVMK ELVERFALYGAIEQYNALDEYPAEDFTEVYLI KFMNLQSARTAKRKMDEQSFFGGLLHVCYA PEFETVEETRKKLQMRKAYVVKTTENKDHY VTKKKLVTEHKDTEDFRQDFHSEMSGFCKA ALNTSAGNSNPYLPYSCELPLCYFSSK

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
1225	2575	A	9710	peptide sequence 1	163	/=possible nucleotide deletion, \=possible nucleotide insertion RSGCVLRMTEWETGAPAVAETPDIKLFGKWS
1006	2556		0710			TDDVHINDISLQDYIAGVRLILL
1226	2576	Α	9713	82	492	QGLPSFLPAFGPSGSWLGPAPTLGSSCNTVDT ICHGYSEIRPLFYLSFCDLLLGLCWLTETLLYG ASVANKDIICYNLQAVGQIFYISSFLYTVNYI WYLYTELRMKHTQSGQSTSPLVIDYTCRVCQ MAFVFSSLI
1227	2577	A	9720	3	416	GKWKRTQVPLLGEECADMDLARKEFLRGNG LAAGKMNISIDLDTNYAELVLNVGRVTLGEN NRKKMKDCQLRKQQNENVSRAVCALLNSGG GVIKAEVENKGYSYKKDGIGLDLENSFSNML PFVPNFLDFMQNGNYF
1228	2578	A	9723	278	411	EASSSNTVASNVADKTDPHSMNSRVFIGNLN TLVLQKSDVEAVF
1229	2579	Α	9725	121	902	LFAMSGFENLNTDFYQTSYSIDDQSQQSYDY GGSGGPYSKQYAGYDYSQQGRFVPPDMMQP QQPYTGQIYQPTQAYTPASPQPFYGNNFEDEP PLLEELGINFDHIWQKTLTVLHPLKVADGSIM NETDLAGPMVFCLAFGATLLLAGKIQFGYVY GISAIGCLGMFCLLNLMSMTGVSFGCVASVL GYCLLPMILLSSFAVIFSLQGMVGIILTAGIIG WCSFSASKIFISALAMEGQQLLVAYPCALLYG VFALISVF
1230	2580	A	9739	11	247	TFVLNMNTPKEEFQDWPIVRIAAHLPDLIVYG HFSPERPFMDYFDGVLMFVDISGKCKRDVCL MWMSNRLAWEFTCRA
1231	2581	A	9744	37		TPLFDFWPGFVLSWLQPLSASLRARRAASGPP ACRIMPTTVDDVLEHGGEFHFFQKQMFFLLA LLSATFAPIYVGIVFLGFTPDHRCRSPGVAELS LRCGWSPAEELNYTVPGPGPAGEASPRQCRR YEVDWNQSTFDCVDPLASLDTNRSRLPLGPC RDGWVYETPGSSIVTEFNLVCANSWMLDLFQ SSVNVGFFIGSMSIGYIADRFGRKLCLLTTVLI NAAAGVLMAISPTYTWMLIFRLIQGLVSKAG WLIGYILITEFVGRRYRRTVGIFYQVAYTVGL LVLAGVAYALPHWRWLQFTVALPNFFFLLY YWCIPESPRWLISQNKNAEAMRIIKHIAKKNG KSLPASL
1232	2582	Α	9753	164	517	PGPGMQGPPPITPTSWSLPPWRAYVAAAVLC YINLLNYMNWFIIAGVLLDIQEVFQISDNHAG LLQTVFVSCLLLSAPVFGYLGDRHSRKATMS FGILLWSGAGLSSSFISPRYSWLF
1233	2583	A	9757	25	419	LPAPWTERVRKSEGLVGTCLGDPMASPRTVT IVALSVALGLFFVFMGTIKLTPRLSKDAYSEM KRAYKSYVRALPLLKKMGINSILLRKSIGALE VACGIVMTLVPGRPKDVANFFLLLLVLAVLF FHQLV
1234	2584	A	9765	71	456	RLELDWGFSLHFLPVAYLCPLSSGFEMNVQP CSRCGYGVYPAEKISCIDQIWHKACFHCEVC KMMLSVNNFVSHQKKPYCHAHNPKNNTFTS VYHTPLNLNVRTFPEAISGIHDQEDGEQCKSV FHWD
1235	2585	Α	9767	52	559	IRSGAMSVDKAELCGSLLTWLQTFHVPSPCA SPQDLSSGLAVAYVLNQIDPSWFNEAWLQGI SEDPGPNWKLKVTSGLLIRGQTGEEMTRDGP ARHMSWVMGRKRDRCLVINHLFIHSSMEYSP CARPGHSARNNTDKNLPHTAIILVTSNTYTTI KINFQAGRSGSCL
1236	2586	Α	9770	352	608	FRGEALTVRFLTKRFIGEYASNFESIYKKHLC

No. of N							
mucleotide seq	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence			hod				
124 2591 A 9849 3 365 1620 173 184 187 1		1					
124			1				1=Isoleucine, K=Lysine, L=Leucine,
anino acid residue of peptide residue of peptide residue of peptide sequence T-Threonine, V-valine, W-Tlyroine, N-Stop codon, /-possible nucleotide deletion, \		uence	١.				
Peptide Sequence	uence			914			
peptide			İ				
						Scquence	
LERKQLNLEITPPCSGTQKAKPSLTSELHWA DGFVTVPDISDRSSTAPAKALI 1237 2587 A 9793 266 315 NILAITYPEPFREFLILDSQSNPKAFALTLCHH QKINFOLIPVSDALTPPLVVCLVSFLTHS RYKPTRPVCTTQFQGCS 1238 2588 A 9802 537 967 ELGAGRSDREAMEALVKEISVEDEAVDKNI FDCNKLAFYRRQKQWLSKKSTYRALLDSVT TDEDSTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPLATIONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPLATIONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPLATIONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPLATIONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPLATIONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPLATIONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPPLATIONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPPLATIONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPPLATIONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPPLATIONIFRIK INDESTRPQIBEASKVPLAEITYGEONIFRIK INDESTRPPLATIONIFRIK INDESTRPPLATIONIFRIK INDESTRUCTIONIFRIK INDESTRUCTIONIFRIK INDESTRPPLATIONIFRIK INDESTRUCTIONIFRIK INDESTRUCTIONIFRIK INDESTRUCTIONIFRIK INDESTRPPLATIONIFRIK INDESTRUCTIONIFRIK INDESTRUCTIONIFRIK INDESTRUCTIONIFRIK INDESTRUCTIONIFRIK INDESTRPPLATIONIFRIK INDESTRUCTIONIFR	l	ł	1		,	ł	
DOFVIVYDISDRSSNAFAKALIT.CHE	\	 	 	 	Sequence		
1237 2387							
QKIKNFQLPVSIDALTPPLVVCFLVSFLTHFS RYKPTRPVCTICQCS	1237	2587	A	9793	266	515	
1238		1	**	****		1 ***	
1238			İ				
FRDCNKIAFYRRGKØWLSKKSTYRALLDSYT TIDEDSTRQININASKYPILABIYGEGNIFIKK NIEETPL KPREVPDVLTSKFSTVRLISCSGDT GSLLADGKGDLKC GSLLADGKGDLKC GSLLADGKGDLKC GSLLADGKGDLKC GSLLADGKGDLKC GSLLADGKGDLKC GSLRADGKGDLKC GSLRADGKGDLKC GSLRADGKGDLKC GSLRADGKGDLKC GSLRADGKGDLKC GSLRADGKGDLKC GSLRADGKGDLKC GSLRADGKGDLKC GSLRADGKGDLKC GSLRADGKGDLGSYK GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGGCCGAGWVABURG GSLRADGKGCGGCCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGGCCGAGWVABURG GSLRADGKGCGGCCGAGWVABURG GSLRADGKGCGGCCGAGWVABURG GSLRADGKGCGGCCGAGWVABURG GSLRADGKGCGGCCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGCGCGAGWVABURG GSLRADGKGCGGACCGGGACCCGAGWVABURG GSLRADGKGCGGACCGGACCCGGACWCABURG GSLRADGKGCGACCGCGGACCCGGACWCABURG GSLRADGKGCGACCGGACCCGGACCCGGACWCABURG GSLRADGKGCGACCGGACCCGGACCCGGACWCABURG GSLRADGKGCGACCGGACCCGACWCABURGCCGACCCGCGACCCGGACCCGGACCCGGACCCGGACCCGGACCCGGACCCGGACCCGACCCGACCCGACCCGACCCGACCCACCCGCACCCACCCGCACCCCACCCACCCCCC	1238	2588	A	9802	537	967	
1239 2589 A 9805 105 540 VPGPPANYRAGAVGAHLPASGIDIFGDLKK SULLADGEGILK SULLADGEG	1222		["	1		
INEETPILRPREVPUNTISKPSTVRLISCSGDT	1		1				
CSS.LLADGKGDLKC	ŀ						1
1239	İ					}	
NNIKRQLYYQVINFAMIYSSALMIWKGLINLT	1239	2589	A	9805	105	540	
1240 2590 A 9819 3 305 TDGRDPLFLATKVEVENDENDLY TDGRDPLPCAARRAGGGGECCGAGWAEWS POPLDPRAMILWQGPVLEAVACQDNDDYLR YGLIFEDLDCNGDGYVDIELQEGLRNWSSAF POPLDPRAMILWQGPVLEAVACQDNDDYLR YGLIFEDLDCNGDGYVDIELQEGLRNWSSAF POPNSEEHG 1241 2591 A 9834 841 1209 SPARGKSNRTDVMITAPKNKKMTENLAAPEA LDSSTHSSSTATQSRAKMNTPAPIPSTYPAIPR GGSGGPPCAPIBNYSSVLQCDTQAMDHKTE SSHSVVEFLFKRTKTTSPFHPAVRSTRN 1242 2592 A 9843 3 589 TISCGPATEPPASLLSSASSDDFCLEKTEDRYS LGSSLDSGMRTPLCRICFQGPEGGELLSPCRC DGSVKCTHQPCLIKWISERGCWSCELCYYKY HVIAISTRNPLQWQAISLTVIEKVQVAAALICS LFILASISWLIWSTFSPSARWQRQDLFQICYG MYGFMDVMVAVDSEDMVQAAKEVGRWS DIPP 1243 2593 A 9846 198 411 WRISHHAGKMPVMKGLLAPQNTFLDTIATRF DGTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPILAYRLIGEEGVAL PANGAGGPGGASARKLSTFLGVVVTVLSMF SIVVELRIGFVGVGHGALQLAMILVAYFILA LTVLSVCAIATNGAVQGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGSIGLMFYLA NVCGCAVSLICULVESVLDVFGA KSKCRFPECLSEGFOFMRKEALSSSQVGEA ANLDEPGGAAGSLTVVYISEHSSLLPQDMM SYIGFKRTAVVKGIMHREAFNIGRIVQVAQ AMSLTEDVIAAALADH.PEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQLLYTVAMLGVNP RFDSASSSYJLDMHSLPFIVNPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKOGAPVAT NAFHSPRWGGMVYNVDSKTYNASVLPVRV EVDMYRVMEVFLAQLRILFGIAQPQLPFKCL LSGPTSEGIMTWELQRLWARSVENLATATT TLTISLA SASGGRPPIAMTVGNYCEAEGPYGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RREPPAGADSLSWAGPRISSYV PVNKKMTRSCSAVGCSTRDTVLSRERGLSF HQPFTDTIQRSKWIRAVNRVDPRSKYWSPUFS GABLCSKHPRESSTEYS (SIRRILKKGAVYSVS SACHSCHESSTERTIKKGAVYSVS SACHSCHRESDFESS (SIRRILKKGAVYSVS SACHSCHRESDFESS (SIRRILKKGAVYSVS SACHSCHRESDFESS (SIRRILKKGAVYSVS SACHSCHRESDFESS (SIRRILKKGAVYSVS SACHSCHRESDFESS (SIRRILKKGAVYSVS SACHSCHRESDFESS (SIRRILKKGAVYSVS SACHSCHRESDFESS (SIRRILKKGAVYSVS SACHSCHRESDFESS (SIRRILKKGAVYSVS SACHSCHRESDFESS (SIRRILKKGAVYSVS SACHSCHRESDFESS (SIRRILKKGAVYSVS SACHSCHRESDFESS (SIRRILKKGAVYSVS SACHSCHRESDFESS (SIRRILKKGAVYSVS SACHSC							
PIRAGEIVVFK VEGRDIPIVIRVIKVHEKDNG DIKFI TKGDINNEGDIBGED DIKFI TKGDINNEGDIBGED DIKFI TKGDINNEGDIBGED DIKFI TKGDINNEGDIBGED DIKFI TKGDINNEGDIBGED DIKFI TKGDINNEGDIBGED DIKFI TKGDINNEGDIBGED DIKFI TKGDINNEGDIBGED DIKFI TKGDINNEGDIBGED DIKFI TKGDINNEGDIBGED DIKFI TKGDIN DIKFI TKGDI DIK							
1240 2590			1	l l			· ·
1240							
PQPLDPAMIL WMQGFVLEAVACQDNDDYLR	1240	2590	A	9819	3	305	
1241 2591							
DPNSEHG							
LDSSTHSSSTATQSRAKMNTPAPTPSTVPAIPR GGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE SHSVVEPLFKRIKTTSPFHPAVRENRN 1242 2592 A 9843 3 589			1			j	
1242 2592 A 9843 3 589	1241	2591	A	9834	841	1209	SPARGKSNRTDVMITAPKNKKMTENLAAPEA
1242 2592 A 9843 3 589							
1242 2592 A 9843 3 589			1	1			
LGSSLDSGMRTPLCRICFGGPEGGELLSPCRC DGSVKCTHQPCLIKWISERGCWSCELCYYKY HVIAISTKNPLQWQAISLTVIEKVQVAAAILGS LFLIASISWLIWSTFSPSARWQRQDLLFQICYG MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DIPP		İ	1			ĺ	
DGSVKCTHOPCLIKWISERGCWSCELCYYKY	1242	2592	A	9843	3	589	TISCGPATEPPASLLSSASSDDFCI:EKTEDRYS
HVIAISTKNPI_QWQAISLTVIEKVQVAAAILGS			1				LGSSLDSGMRTPLCRICFQGPEQGELLSPCRC
LFLLASISWLIWSTESPSARWQRQDLLFQICYG MYGFMDVMIVAVDSEDMVQAAKEVGKRWS DIPP				1			DGSVKCTHQPCLIKWISERGCWSCELCYYKY
1243 2593 A 9846 198 411 WRISHHAGKMPVMKGLLAPQNTFLDTIATRF DGTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLIGEEGVAL PANGAGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMILVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMILDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDK WSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLW ARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVEPRHPDVRAPLRVTSPGLRSS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAYGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	ŀ]	1				
1243 2593 A 9846 198 411 WRISHHAGKMPVMKGLLAPQNTFLDTIATRF DGTHSNFILANAQVAKGPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMILVAYPILA LTVLSVCAIATNGAVQGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFFEGLSEGFGPMRKEALSSGSVQEAE AMI.DEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGFKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVI.AAALADHI.PEDK WSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQIRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRILLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTILVPSTRMALGTLALVLALPCK RERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAYGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS]	1		t 		
1243 2593 A 9846 198 411 WRISHHAGKMPVMKGLLAPQNTFLDTIATRF DGTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMILVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGIMFYLA NVCGCAVSLLGLVESVLDVFGA 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGFKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRPRFIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK REERPAGADSLSWGAGPRISSSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQPFTDTIQRSK WIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRKLKKGAVPSVS							
DGTHSNFILANAQVAKGFPIVYCSDGFCELAG FARTEVMQ 1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGFEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGFMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVNPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQCFTDTIQRSKWIRAVNRVDPRSKKIWIFGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							1
1244 2594 A 9848 116 650 PICGFLYLCSAMASESSPLLAYRLLGEEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMILVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGFEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1243	2593	A	9846	198	411	
1244 2594 A 9848 116 650 PICGFLYLČSAMASESSPLLAYRLLGEEGVAL PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMILLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGFTSLEA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
PANGAGGPGGASARKLSTFLGVVVPTVLSMF SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDK WSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
SIVVFLRIGFVVGHAGLLQALAMILLVAYFILA LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1244	2594	A	9848	116	650	
LTVLSVCAIATNGAVQGGGAYCILQHRWTG VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDK WSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS		i	1	}	}		
VWPVLPAREVMISRTLGPEVGGSIGLMFYLA NVCGCAVSLLGLVESVLDVFGA 1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFFTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDK WSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS				1	ļ		
1245 2595 A 9849 573 1620 KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDK WSAEKRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS						1	
AMILDEPQEQAEGSLTVYVISEHSSLLPQDMM SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1245	2505	 	0040	572	1620	
SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1245	2393	A	9849	3/3	1020	
AMSLTEDVLAAALADHLPEDKWSAEKRRPL KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS				1			
KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS		1	l	1			
YVQPFLNALGAAGNFSVDSQILYYAMLGVNP RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
RFDSASSSYYLDMHSLPHVINPVESRLGSSAA SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSK WIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS		1	1				
NAFHSPRWGGIMVYNVDSKTYNASVLPVRV EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSK WIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS		1]			
EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSK WIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS		1	ŀ				
LSGPTSEGLMTWELDRLLWARSVENLATATT TLTSLA 1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSK WIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS					l		
1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSK WIRAVNRVDPRSKKJWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS			1		1		,
1246 2596 A 9850 114 464 PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSK WIRAVNR VDPRSKKJWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	ĺ		Í	1		1	
ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1246	2596	A	9850	114	464	
QDGLSPCFFFTLVPSTRMALGTLALVLALPCK RRERPAGADSLSWGAGPRISSYV 1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	0		l				
1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS]				
1247 2597 A 9851 2 327 FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
HQFPTDTIQRSKWIRAVNRVDPRSKKJWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS	1247	2597	A	9851	2	327	FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF
GAILCSKHFQESDFESYGIRRKLKKGAVPSVS							
		1]	1			
		1	l	1			LYKVFKYSSRCTS

SEQ ID NO: of nucl- eotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first amino acid residue of peptide sequence	to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1248	2598	A	9853	58	444	RVDDFVYSKGGKDAGGADVSLACRRQSIPEE FRGITVVELIKKEGSTLGLTISGGTDKDGKPR VSNLRPGGLAARSDLLNIGDYIRSVNGIHLTR LRHDEIITLLKNVGERVVLEVEYELPPPGGCP WT
1249	2599	A	9856	2	1265	LPPPRPSRHRRGRAGTRASAAAAAGPTVSAV RAPVRGQDSGAGTPQGRLAGRGAHLSRVGA SGSGVAAGPAARHAPRRCADAGEAVGASC GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT PMGAGDAGASAESAVTTAPQEPPARPLQAGS GAGPAPGRAMRSTTLLALLALVLLYLVSGAL VFRALEQPHEQQAQRELGEVREKFLRAHPCV SDQELGLLIKEVADALGGGADPETNSTSNSSH SAWDLGSAFFFSGTIITTIGGGGDWHVGGGK ELPHGGRCRETEGSQVAPRLPASPLCPGYGN VALRTDAGRLFCIFYALVGIPLFGILLAGVGD RLGSSLRHGIGHIEAIFLKWHVPPELVRVLSA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVIVTLTTVGFGDYVA
1250	2600	Α	9873	2	652	FVVPSPCGGIPGRAPNGASRPTMGNSASRNDF EWVYTDQPHTQRRKEILAKYPAIKALMRPDP RLKWAVLVLVLVQMLACWLVRGLAWRWLL FWAYAFGGCVNHSLTLAIHDISHNAAFGTGR AARNRWLAVFANLPEGVPYAASFKKYHVDH HRYLGGDGLDVDVPTRLEGWFFCTPARKLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV OLA
1251	2601	Α	9875	150	1209	PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR LESYRPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPL SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDICPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS
1252	2602	A	9879	6	376	KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ
1253	2603	Α	9880	180	388	KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC
1254	2604	A	9881	19	494	VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATTFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL
1255	2605	A	9896	72	386	RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGQTT
1256	2606	A	9902	95	399	SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE
1257	2607	Α	9905	374	459	EHLKSTPNRLGVVAHTCNPSTLGGRGGW

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ĺ	ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
}		ļ	j	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
[}		residue of peptide	sequence.	/=possible nucleotide deletion, \=possible
}			ļ	sequence		nucleotide insertion
1258	2608	A	9911	364	1974	AGPGVPAVGGRWASGPGLGGRTLCSGPPDH
1230	2000		,,,,,	1 301	***	QRRGPSCGASGDPQCVGSPHPQRARPLLARP
	ļ			}	J	GARLLPGHLPSPRPPRLPTGQPPAAAFRGPVR
			ļ		}	PQGGGHIHPLPTPGGRPCFAVSEGSGSALLLS
						YLGECGSSSYVTGAACISPVLRCREWFEAGLP
		1	1		1	WPYERGFLLHQKIALSRYATALEDTVDTSRL
}	}	}]		ļ	FRSRSLREFEEALFCHTKSFPISWDAYWDRND
	l	1				PLRDVDEAAVPVLCICSADDPVCGPPDHTLTT
1	ļ	l		j		ELFHSNPYFFLLLSRHGGHCGFLRQEPLPAWS
				l		HEVILESFRALTEFFRTEERIKGLSRHRASFLG GRRRGGALORREVSSSSNLEEIFNWKRSYTRL
1		1	Ì	}	1	MAAAAGAAAAPGSREPQDRPECGAGHPGPR
			ł	ļ		YYRHPERWLLRPEAFLGPLRTRAPSAEDSQR
		1		İ		ERPAARSGPEMRVRYPVVAAVLAPYLALSQD
j]	Į	}	1		PMVKSSASGQGASGSYNHVREEMLIKAGGA
		İ				MSRRVVRQSKFRHVFGQAAKADQAYEDIRV
Į.		l	ŀ	}		SKVTWDSSFCAVNPKFLAIIVEAGGGGAFIVL
	<u></u>					PLAK
1259	2609	A	9919	693	935	GCFKFIGESTCCWIFPSSVTTQCVVAKAPRAA
		1		1		TLSKAERLRSQPGPEQGGSSYRPRTPTAAAIL
10/0	12610	 	0001	455	1082	PPRPGRSHRKRKLVSTK QRSCLCSAIEKDGGDVKALYRRSQALEKLGR
1260	2610	A	9921	455	1082	LDQAVLDLQRCVSLEPKNKVFQEALRNIGGQ
			l	İ	1	IQEKVRYMSSTDAKVEQMFQILLDPEEFGTE
				,		KKQKASQNLVVLAREDAGAEKIFRSNGVQLL
	Ì			1		QRLLDMGETDLMLAALRTLVGICSEHQSRTV
		ł	ł	ł	1	ATLSILGTRRVVSILGVESQAVSLAACHLLQV
	<u> </u>	<u> </u>				MFDALKEGVKKGFRGKEGAIIV
1261	2611	A	9928	1	438	GFRGAEAPGAAQAPKKKKPRPTEGGPGAGSG
		ļ	1	1		RGKDPYRGPTLLHQPKPPKDEFLSSLESYEIAF
		Ī	ĺ	[[PTRVDHNGALLAFSPPPPQRQRRGTGATAES RLFYKEASPSTHFLLNLTRSSRLLAGHVSVEY
			1			WTREGLAWQRADRPHCLYA
1262	2612	A	9931	168	435	AAEMGRAGAAAVIPGLALLWAVGLGGPPPA
,,,,,,	5015	**		1 200	1.55	PPRLPFCLQELQGRHALHTFSLERTCSYQDFL
		1	1			WADEGRLLHVGAQDLATWHTLSPLGLW
1263	2613	A	9938	247	488	RMSATSVDQRPKGQGNKVSVQNGSIHQKDG
			1	-		CNDDDFEPYLRSPDNQSNSYPPMSDPYMPGY
		<u> </u>	L	<u> </u>	L	YAPSIGFPYSLGEAAWSQL
1264	2614	A	9941	61	277	ESIGLTALGPRRPWEHRWSDPITLKMKGWG
				}	1	WLALLLGALLGTAWARRSQDLHCGACKAVR
1265	2615	A	9956	2	522	RRVRQFNIYDY FVASEVSKMPVPASWPHPPGPFLLLTLLGLT
1203	2013	A	7730	*	ا عدد	EVAGEEELQMIQPEKLLLVTVGKTATLHCTV
}	1		1		1	TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP
						RVTTVSDLTKRNNMDFSIRISSITPADVGTYY
	ļ	1	l	ł		CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG
						FLSQVWWWLSSHPFMN
1266	2616	A	10002	243	387	PKNNACHLLFTAVCQPRCKHGECIGPNKCKC
		<u> </u>	1		L	HPGYAGKTCNQGRKTV
1267	2617	A	10004	36	707	LPAPASTWSVARETMASSSVPPATVSAATAG
1		1	1	1		PGPGFGFASKTKKKHFVQQKVKVFRAADPLV
1		1		I		GVFLWGVAHSINELSQVPPPVMLLPDDFKAS
	1			1	l	SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS
	1	1				YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS
		1	1	Ī		SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP
1268	2618	A	10005	2	209	GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP
		[1			SQDELEHSLGESAAQGAAGVVLWVSWENTR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino gaid common (A-1)
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1.00	in NO.	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	иелсе		09/496	соттевроле	to last amino	1-Isoleucine, K=Lysine, L=Leucine,
uence	Beriet	ľ	914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline,
delice		1	717	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine,
		ŀ	1	residue of		T=Threonine, V=Valine, W=Tryptophan,
ľ			ŀ	peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			}			/=possible nucleotide deletion, \=possible
<u> </u>			<u> </u>	sequence		nucleotide insertion
1260	2619		10010	100	600	TKVSLGLA
1269	2619	A	10010	245	688	FGMLKNKGHSSKKDNLAVNAVALQDHILHD
1		ĺ				LQLRNLSVADHSKTQVQKKENKSLKRDTKAI
	1	1				IDTGLKKTTQCPKLEDSEKEYVLDPKPPPLTL
1						AQKLGLIGPPPPPLSSDEWEKVKQRSLLQGDS
						VQPCPICKEEFELRPQVFSIRG
1270	2620	A	10011	2	588	RVDDFVRPLPPGLMSRSRASIHRGSIPAMSYA
(1			20	PFRDVRGPSTHRTQYVHSPYDRPGWNPRFCII
1	1	ļ				SGNQLLMLDEDEIHPLLIRDRRSESSRNKLLR
ļ						RTVSVPVEGRPHGEHEYHLGRSRRKSVPGGK
1 .	l	i				QYSMEGAPAAPFRPSQGFLSRRLKSSIKRTKS
1		1				QPKLDRTSSFRQILPRFRSADHDRYRGWSMW
!						DEIDV
1271	2621	A	10013	209	363	LPAPPNLSPRLSFGFQFPGGNDNYLTITGPSHP
						FLSGAEVSQSCRRRGGRA
1272	2622	A	10014	7	388	SAVTISWKWRSVMGIQTSPALLASLGAGLVT
) :	1	1	ļ			LLGLAVGSYLVRRSRRPQVTLLDPNEKDLLR
						LIDKTLSARSPCKHIYLSTRIDGSLSIRPYTPVT
ł	ł	i				SDEDQGYVDIDIKVYLKGVHPTFPEGGKMSH
1273	2623	A	10016	1	1339	MAARTLGRGVGRLLGSLRGLSGQPARPPCGV
		1	10010	•	1557	SAPRRAASGPSGSAPAVAAAAAQPGSYPALS
	1	1	1			AQAAREPAAFWGPLARDTLVWDTPYHTVW
•						DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS
	1	1			,	PESVALIWERDEPGTEVRITYRELLETTCRLA
j 1]]				NTLKRHGVHRGDRVAIYMPVSPLAVAAMLA
		ĺ				CARIGAVHTVIFAGFSAESLAGRINDAKCKVV
1	1	Į.				ITFNQGLRGGRVVELKKIVDEAVKHCPTVQH
						VLVAHRTDNKVHMGDLDVPLEQEMAKEDP
		[[
) !		}				VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT
ļ i		ŀ				QAGYLLYAALTHKLVFDHQPGDIFGCVADIG
						WITGHSYVVYGPLCNGATSVLFESTPVYPNA
						GRYWETVERLKINQFYGAPTAVRLLLKYGD
ĺ						AWVKKYDRSSLRTLGSVGEPINCEAWEWLH
1274	2624		10017		2550	RVVGDSRCTLVDTWWQT
12/4	2024	A	10017	1	3750	FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG
						KTLGSFFGSLPGFSSARNLVANAHSSARARPA
						ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE
						KELQPSEKMVSGAKDLVCSKMSRAKDAVSS
1		ł	1			
		i			1	GVASVVDVAKGVVQGGLDTTRSALTGTKEV
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG
i I				:		VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV
				:		VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSWVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKDTVCSGVTGAVNV LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKDTVCSGVTGAVN LAKEAIQGGLDTTKSWVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTK TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT
					· Y	VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKDTVCSGVTGAVN LAKEAIQGGLDTTKSWVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTK TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV LTGTKDTVTTGVMGAVNLAKGTVQTGVETS KAVLTGTKDAVSTGLTGAVNVARGSIQTGV DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA KGTIQTGVDTTKTVLTGTKDTVCSGVTGAVN LAKEAIQGGLDTTKSWVMGTKDTMSTGLTG AANVAKGAMQTGLNTTQNIATGTKDTVCSG VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC SGVTGAANVAKGAVQGGLDTTKSVLTGTKD AVSTGLTGAVNVAKGTVQTGVDTTK TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV VIGTKDTMSTGLTGAANVAKGAVQTGVDTA KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM DTTKTVLTGTKDTTYSGVTSAVNVAKGAVQT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid	Predicted end nucleotide location corresponding to last amino acid residue of peptide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				orquenes .		VAKGAIQGGLDTTKSVLTGTKDAVSTGLTGA VKLAKGTVQTGMDTTKTVLTGTKDAVCSGV TGAANVAKGAVQMGVDTAKTVLTGTKDTV CSGVTGAANVAKGAVQTGLKTTQNIATGTK NTLGSGVTGAAKVAKGAVQGGLDTTKSVLT GTKDAVSTGLTGAVNLAKGTVQTGVDTSKT VLTGTKDTVCSGVTGAVNVAKGTVQTGVDT AKTVLSGAKDAVTTGVTGAVNVAKGTVQTG VDASKAVLMGTKDTVFSGVTGAMSMAKGA VQGGLDTTKTVLTGTKDAVSAGLMGSGNVA TGATHTGLSTFQNWLPSTPATSWGGLTSSRT TDNGGEQTALSPQEAPFSGISTPPDVLSVGPEP
						AWEAAATTKGLATDVATFTQGAAPGREDTG LLATTHGPEEAPRLAMLQNELEGLGDIFHPM NAEEQAQLAASQPGPKVLSAEQGSYFVRLGD LGPSFRQRAFEHAVSHLQHGQFQARDTLAQL ODCFRL
1275	2625	A	10025	124	415	TILARKKEKTCPCKKEIGRNSRSGMYSRKAM YKRKYSAANTKVEKKKKEKVLAPVTKPVGG DKNGGTRVVKLPTMPRYYPTEDVPRKLLSHG KKPFS
1276	2626	A	10030	3	507	GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSPMS ASRPQSPTTPWCLPRRYMKHKRDDGPEKQED EAVDVTPVMTCVFVVMCCSMLVLLYYFYD) LVYVVIGIFCLASATGLYSCLAPCVRRLPFGK CRIPNNSLPYFHKRPQARMLLLALFCVAVSV VWGVFRNEDQ
1277	2627	A	10035	51	869	YSRFTVPLPATMASSEVARHLLFQSHMATKT TCMSSQGSDDEQIKRENIRSLTMSGHVGFESL PDQLVNRSIQQGFCFNILCVGETGIGKSTLIDT LFNTNFEDYESSHFCPNVKLKAQTYELQESN VQLKLTIVNTVGFGDQINKEERQLGRSQSTEN PQKYRSEQHPVEPKKCTSFWKGALGKWAGIE SSGQSAQQPYLPINSPPHRLADVADVHLFSSV LSGAFGCYHLDVTVNEFKKQQNRDEQEGYS KGDQEQGSWKHGADPLRGGEM
1278	2628	A	10036	3	457	RAFDVRRKKSLRPCCPRDFHAGCLTVSGPST VMGAVGESLSVQCRYEEKYKTFNKYWCRQP CLPIWHEMVETGGSEGVVRSDQVIITDHPGDL TFTVTLENLTADDAGKYRCGIATILQEDGLSG FLPDPFFQVQVLVSSASSTENSVKTP
1279	2629	A	10039	214	435	NDSLVPMSSWRSCARAPSSESAWRRSAATRR SRKCLRTKRKRWSSGKGTQMQSTLSETPRRA QMPCMWWYPFWG
1280	2630	A	10043	2	344	RATWHNAGKEREAVQLMAGAEKRVKASHS FLRGLFGGNTRIEEACEMYTRAANMFKMAK NWSAAGNAFCQAAKLHMQLQSKHDSATSFV DAGNAYKKADPQGKTARHVACYLCV
1281	2631	A	10080	620	818	VIYKLDSSLFSYFIYFFIFETESHFLPLMKWTG PIMAHCSLKILASRNSADSAFLSAGDTSLSHST
1282	2632	A	10084	3	1640	SASIIIRGDKRASGEVGIAPSSRHILIGEPSAKY NGTAIISLVRGPGILGEVTVFWRIFPPSVGEFA ETSGKLTMRDEQSAVIVVIQALNDDIPEEKSF YEFQLTAVSEGGVLSESSSTANITVVASDSPY GRFAFSHEQLRVSEAQRVNITIIRSSGDFGHVR LWYKTMSGTAEAGLDFVPAAGELLFEAGEM
						RKSLHVEILDDDYPEGPEEFSLTITKVELQGR GYDFTIQENGLQIDQPPEIGNISIVRIIIMKNDN AEGIIEFDPKYTAFEVEEDVGLIMIPVVRLHGT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YGYVTADFISQSSSASPGGVDYILHGSTVTFQ HGQNLSFINISIIDDNESEFEEPIEILLTGATGG AVLGRHLVSRIIIAKSDSPFGVIRFLNQSKISIA NPNSTMILSLVLERTGGLLGEIQVNWETVGPN SQEALLPQNRDIADPVSGLFYFGEGEGGVRTII LTTYPHEEIEVEETFIIKLHLVKGEAKLDSRAK DVTLTIQEFGDPNGVVQFAPETLSKKTYSEPL
1283	2633	A	10088	316	516	ALEGPLLITFFVRRVKGTFGEIM MGSKTLPAPVPIHPSLQLTNYSFLQAVNGLPT VPSDHLPNLYGFSALHAVHLHQWTLGYPAM HLXRS
1284	2634	A	10091	2	569	FVSPSRAMASALIYVSKFKSFVILVVTPLLLLP LVILMPAKFVRCAYVIILMAIYWCTEVIPLAV TSLMPVLLFPLFQILDSRQVCVQYMKDTNML FLGGLIVAVAVERWNLHKRIALRTLLWVGA KPARLMLGFMGVTALLSMWISNTATTAMMV PIVEAILQQMEATSAATEAGLELVDKGKAKE LP
1285	2635	A	10092	290	728	KQSTRPDVMTLYPLHWQEEMSGESVVSSAVP AAATRTTSFKGTSPSSKYVKLNVGGALYYTT MQTLTKQDTMLKAMFSGRMEVLTDSEGWIL IDRCGKHFGTILNYLRDGAVPLPESRREIEELL AEAKYYLVQGLVEECQAALQV
1286	2636	A	10100	1	574	RPRGRGAWAGPGGDYSGVRRQQRRRTRISGS QRGSDAAGTMGCCTGRCSLICLCALQLVSAL ERQIFDFLGFQWAPILGNFLHIIVVILGLFGTIQ YRPRYIMVYTVWTALWVTWNVFIICFYLEVG GLSKDTDLMTFNISVHRSWWREHGPGCVRR VLPPSAHGMMDDYTYVSVTGCIVDFQYLEVI HSA
1287	2637	A	10103	252	376	RSRMGDKPIWEQIGSSFIQHYYQLFDNDRTQL GAIYVSFQL
1288	2638	A	10107	1	478	MEEEDESRGKTEESGEDRGDGPPDRDPTLSPS AFILRAIQQAVGSSLQGDLPNDKDGSRCHGL RWRRCRSPRSEPRSQESGGTDTATVLDMATD SFLAGLVSVLDPPDTWVPSRLDLRPGESEDM LELVAEVRIGDRDPIPLPVPSLLPRLRAWRTG KT
1289	2639	A	10113	237	438	LLSRMPSTNRAGSLKDPEIAELFFKEDPEKLFT DLREIGHGSFGAAYFARDVRTNEVVAIKKMS YSG
1290	2640	Α	10114	367	856	RGAKAKSAVLPPGPPCSSILILSPPAPLTPRSPG TEATRPTAMSKSLKKKSHWTSKVHESVIGRN PEGQLGFELKGGAENGQFPYLGEVKPGKVAY ESGSKLVSEELLLEVNETPVAGLTIRDVLAVI KHCKDPLRLKCVKQGESSGLLSVLPGGGTAR GAGQ
1291	2641	A	10116	128	591	RTIRETERRSALSCSVLKSEPLPGLQPQASQQR RRRLPGRRQVQVQEGGGSGLRAWVLAMASV LGSGRGSGGLSSQLKCKSKRRRRRSKRKDK VSILSTFLAPFKHLSPGITNTEDDDTLSTSSAE VKENRNVGNLAARPPPSGDRARGGATR
1292	2642	A	10121	1	749	QRRFFRAGLWGGHGLTDGLRRNGGCGCSAR VPRVGERLRGHRCPDPLCLLLDMLFLSFHAG SWESWCCCCLIPADRPWDRGQHWQLEMADT RSVHETRFEAAVKVIQSLPKNGSFQPTNEMM LKFYSFYKQATEGPCKLSRPGFWDPIGRYKW DAWSSLGDMTKEEAMIAYVEEMKKIIETMP MTEKVEELLRVIGPFYEIVEDKKSGRSSDITSD

No. of No. of	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide sequence where the corresponding to corresponding to corresponding to the correspo							
Docation Docation			1"00				
uence unce 09/496 correspondit on the first of the state of the state of the state of the state of peptide residue of peptide sequence common and the state of peptide sequence common and th			j				
United Part United Part United Part P							
mino acid recidue of peptide sequence peptide sequence se	1 •	ucirce	ļ	1			
residue of peptide sequence	delice		ĺ	714			
Peptide			ĺ				
	1		[Sequence	
LGNVLTSTFNAKTVNGKAESSDSGAESEEEE			ĺ				
1293			ļ <u></u> .	 	sequence		
1293	1				ļ·		
	1202	2642		10124	<u> </u>	000	
CGRGIRDSARMCSTCACVEYYGKALEFUL, GVKINCHIGKMKYRNIKIMERPERLQSGILV CTDVMARGIDIPEVNWVLQYDPPSNASAY (GVKINCHIGKMKYRNIKIMERPERLQSGILV CTDVMARGIDIPEVNWVLQYDPSNASAY (AVCRINGIBER SYNFLAIN QKCPLQEMKPQRNTADLLPKLKSMALADRA VPEKGMKARVSYYQAYAKHECNILFRLKDL) DPASIARGFALIRMYKMPELRIKOGPTDFVPV DVNTDTIPFKDKIREKQRQKILEQQRREKTEN EGRKFIKINKAWSQKAKKK. 1294 2644 A 10129 91 1042 VTMYKDCIESTIGDYFLLCDAEGFWGILLEQQRREKTEN EGRKFIKINKAWSQKAKKK.	1293	2043	A	10124	2	989	
GVKIMCHIGKMKYKRNIKIPMERRKLÖSGLV CTDVWARGDIPPEVNWU_QUPPSNASAPVH RCGRTARIGIGGSAL VFILIPMEBSYNFILAN QKCPLQEMKPQRNTADLLPKLESYNFILAN QKCPLQEMKPQRNTADLLPKLKSMALADRA VFEKGMKAPVSYVQAYAKHECHLIFRLKDL DFASLARGFALLRMPKMPELRGKGFIDEPVD DVNTDTIPFKDKIERKGRGKLLEQGREKKTEN EGRRYFIKNKAWSKQKAKKK 1294 2644 A 10129 91 1042 VTMYKDCIESTIGFYFILCDAEGFWGIILESLA LIGIVVTILLILAFLFLMRKIQDCSQWNULPY LIFLISVIGIFGLAFAFIIENQOTAPVRYFLF GVLFALCTSCLLAHASNLVKLVRGCVSFWW TUCIAIGCSLLQIMATEVYTLIMTGMMFWN MTPCQLNVDFVVLLVYYLFLMALTFFVSKA TUCIAIGCSLLQIMATEVYTLIMTGMMFWN MTPCQLNVDFVVLLVYYLFLMALTFFVSKA TUCIAIGCSLLQIMATEVYTLIMTGMMFWN MTPCQLNVDFVVLLVYYLFLMALTFFVSKA TUCIAITCAGSLLQIMATEVYTLIMTGMMFWN MTPCQLNVDFVVLLVYYLFLMALTFFVSKA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA HSRQVPNQELSRNKWVLLINSDFLSHSGA 1295 2645 A 10133 376 518 RPRSVTHMSQVPT TUCKTTCCRPSCCVSSCCRPQCQSVCQPT CSRPSCQTTTCCRTTCVPSCCVSSCCRPQCC 1296 2647 A 10138 48 407 MVSSCCGSVCSDQGCGQDLQETCCRPSCCE 1297 2647 A 10138 48 407 MVSSCCGSVCSDQGCGQDLQETCCRPSCCE 1298 2648 A 10156 94 453 GGRKSAEMSSQVPTTPASGCYTNISMTTEG GRMRKSAEMSSQVPTTPASGCYTNISMTTEG CRUPTODLYNCLOSPEC 1299 2649 A 10161 1 393 PRSSLVDGGGRVSAERGEGEGEGEGEGEGEGEGEGEGEGEGEGEGEGEGEGEG		l	[
CTDVMARGIDPEVNWVLQYDPPSNASAP			1	ł		ļ	ſ
RCGRTARIGHGGSALVFLLPMEESVIRLAIN							1 ,
QKCPLQEMKPQRNTADLLPKLKSMALADRA	i		İ		İ		1
	-		l				
DFASILARGÉNLIRUÉRMERGICG/PEDFVPV	1		ĺ				
1294 2644 A 10129 91 1042 VTMYKDCIESTGDYFLLCDAEGPWGIILESLA ILGIVVTILLILAFI,FLMRKIQDESGWNVLFTQ LLFLLSVLGHGLAFFILENQQTAFVRYFLF GVLFALCFSCLLAHASNLVKLVRGQKSYSWT TILCIAIGGSLA,QIMATEYVTILMTRGMMFVN MTPCQLNVDFVVLLVYVLF,MALTFFVSKAT FCGPCENWKQHGRL,BITTSJIWVVWISML LRGNPQFQRQWDDPVVCIALVTNAWVFL LLYIVPELCILYRSCRQECPLQONACPVTAYQ HSPQVENQELSRDKWKVLINSDFLSHSGA FRQVENQELSRDKWKVLINSDFLSHSGA GRGAPRQEJSRDKWKVLINSDFLSHSGA GRGAPRQEJSRDKWKVLINSDFLSHSGA GRGAPRQEJSRDKWKVLINSDFLSHSGA GRGAPRQEJSRDKWKVLINSDFLSHSGA FRVENVTHNSQWCFLFQDHFWLFQSGAPG GRGAPRQEJSRDKWKVLINSDFLSHSGA GRGAPRQEGTGSWRQV 1296 2646 A 10135 3 551 EWSLDFFMGIMSGQVGDLSSPGGSLAQFRE NIQQVI SALPPDDYPTLJRWLQARSFDLQK SASKQELLABPRDSYPCLLJRWLQARSFDLQK SASKQELLABPSRSCELLJRECELGSQKLGKR VEKINAFGLEGIGLRDLWKPGELLQR VNANGIGGHDGEGSPVWYHIVGSQGPFGLLQR GVRSCCSQCGQCQDLCQETCCRPSCCE TTCCRTTCCRPSCCVSSCCRPQCCQDVCCQPT GSRSSCQCTTCCRTTCYRSCCVSSCCRPQCCQDVCCQPT GSRSSCQCTTCCRTTCYRSCCVSSCCRPQCCQDVCCQPT GSRSSCQCTTCCRTTCYRSCCVSSCCRPQCCQPVCCQPTCCRPSCCETTCCHPXCC GGRNSCAEMFSQVPRTFASGCYYLISMTFEQ GGMYLFFDQTTRRSPTWSKILARRQLVTKI QQEMALFRAGGPQYAQLYEDSQFP RIVAVKNDHDFLEDLGDEFPLCRRINT A 10161 1 393 PRFSELVDGRRVSARFGGSPSKAATVRSQFT A 10161 1 393 PRFSELVDGRRVSARFGGSPSKAATVRSQFT A 10162 98 391 AKIASLERIMPANYTCTRPDGDMTDFYFTY A VTYTGILGFGLGJGHGHTGT KQLQCMPMEGRGRASSSISDLQCRFCKTG KQLQCMPMEGRGRASSSISDLQCRFCKATG KQLQCMPMEGRGRASSSISDLQCRFCKATG KQLQCMPMEGRGRASSSISDLQCRFCKATG KQLQCMPMEGRGRASSSISDLQCRFCKATG KQLQCMPMEGRGRASSSISDLQCRSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHSTP PSRPLPQAGRCLVAPLRPHPDWXAKTLAKA LRAPGKPWLAAPSPLGDLGAPGLAPGPSTAP RTLSVEFPQUENQLTVAPLTPHTPUCAKTALAE LRAPGKPWLAAPSPLGDLGAPGLAPGPSTAP RTLSVEFPQUENQLTVAPLTPHTPUCAKTALSEPSKP RRLSVEFPQUENQLTVAPLTPHTPUCAKTALSEPSKP RRLSVEFPQUENQLTVAPLTPHTPUCAKTALSEPSKP RRLSVEFPQUENQLTVAPLTPHTPUCAKTALSEPSKP RRLSVEFPQUENQLTVAPLTPHTPUCAKTALSEPSKP RRLSVEFPQUENQLTVAPVTPTVLCLGQK RVLGRPDFGRATCRSSTERGERGREGTGT RTLSVEFPQUENQLTVAPVTPTVLCLGQK RVL	ı	ļ	ł			}	
EGRRFIKNKAWSKOKÄKKK 1294 2644 A 10129 91 1042 VTMYKDCIESTGDYFLLCDAEGPWGIILESLA ILGIVVTILLLAFLFLMRKIQDCSQWNVLPTQ LLFLLSVLGLFGLAFAFILEINQQTAPKYFLF GVLFALGFSCLLAHASIN/KLVRGCVSSSWT TILCIAIGCSLLQIIIATEYVTLIMTRGMFVN MTPCQLNVDFVVLLVVPLMALTFFVSKAT FCGPCBWKQFIGRLIFITVLFSIIWVVWSMML LRGNPQFQRQWDDPVVLLAVTNAWVFL LLYIVPELCILYRSCRGECFLQGNACFVTAYQ HSFQVENQELSRDKWKVLLNSDFLSISGA GRGAPRQEGROSSWRQV GRGAPRQEGROSSWRQV GRGAPRQEGROSSWRQV GRGAPRQEGROSSWRQV GRGAPRQEGROSSWRQV GRGAPRQEGROSSWRQV SEWSLDFFRGMISGGVQDLSPSQEKSLAQFRE NIQDVLSALPNPDDYFLLRWLQARSFDLQKS EDMLRKHMEFRKQQDLANILAWQPPEVVRL YNANGICGHGGGSPWYHOGSQDFKGLLL SASKQELLRDSFRSCELLLRECELQSKLGKR VEKILIFFIGEGLGLRDLWFGSGLGLKG VEKILIFFIGEGLGLRDLWFGSGLGVCGPT CSRPSCQTTCCRTTCYRSCCVSSCCRPQCC GVPVCQPTCCRTSCCVSSCCRPQCC GVPVCQPTCCRTSCCVSSCCRPQCC QPVCCQPTCCRTSCCSTCTCCFPSCC GGNRKSABMFSQVPTRTSPYRMSRLLARRQLVTKI QQETAKEACDWLRAAGSCYYLNSMTPEG GEMYLRFDQTTRRSPYRMSRLARRQLVTKI QQETAKEACDWLRAAGSCYYLNSMTPEG GEMYLRFDQTTRRSPYRMSRLARRQLVTKI QQETAKEACDWLRAAGSCYYLNSMTPEG GEMYLRFDQTTRRSPYRMSRLARRQLVTKI QQETAKEACDWLRAAGSCYYLNSMTPEG GEMYLRFDQTTRRSPYRMSRLARRQLVTKI QQETAKEACDWLRAAGSCYYLNSMTPEG GEMYLRFDQTTRRSPYRMSRLARRQLVTKI QQETAKEACDWLRAAGSCYYLNSMTPEG GEMYLRFDQTTRRSPYRMSRLARRQLVTKI QQETAKEACDWLRAAGSCYYLNSMTPEG GEMYLRFDQTTRRSPYRMSRLARATVRSQFT NIVAVKNDHDFLEKDLGEFLCRRLNT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMKEAVRYQFT STAQLEFMCSKLATSSHOTE KQLQCMPMEGAGRASSISIOLAWAAR LLCSGANRCKVLVRQNSTPNTQQPAVMPSTP PSRPLPQAGRCLVAPLRPHDDWAAKTLAKA LRAPGKPWLAAPSPIQEDLGAARTASSPHOTE SKHYPGVGSARHSPQASASAGGSPWQGKAQT RWLGKPDPGRRKRRGSPGEGGGRNSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVMPSTP PSRPLPQAGRCLVAPLRPHDDWAAKTLAKA LRAPGKPWLAAPSPIQEDLGAARTASSPHOTE SKHYPGVGSARHSPQASASAGGSPWQGERGAA RTLSVE	1		!				
1294	ļ		Ì		ĺ		
ILGIVVTILLIAF,EMRKIQDCSQWNVLPTQ LLFILSVLGIFGLAFAPIEMQCTAPYRYEF GVLFALCFSCLLAHASNLVKLVRGCVSFSWT TILCIAIGCSLQIIIATEYYTLIMRGMFWN MTPCQLNVPFVVLLVYVLFLMALTFYSKAT FCGPCENWKQHGRLIFITVLFSIWVVWISML LRGNPQFQRQPQWDDPVVCIALVTNAWVFL LLYIVPELCILYRSCRQECPLQGNACPVTAYQ HSFQVENQELSRDKWKVLINSDFLSHSGA GRAPRQEGFGSSWRQV HSPQVENGESRDKWKVLINSDFLSHSGA GRAPRQEGFGSSWRQV GRAPRQEGFGSSWRQV EWSLDPMGIMSGQVGDLSPSQEKSLAQFRE MQDVLSALPMPDDYFLLRWLQARSFDLQKS EDMLRKHMEFRKQQDLANILAWQPPEVVRL YNANGICGHDGEGSPVW7HVGSQDPKGLLL SASKQELLRDSFRSCELLLRECELQSQKLGKR KEKIAIFGLEGLGLRDWKPGIELLQE 1297							
LLFILSVLGLFGLAFAFIIELNQQTAPVRYFLF GVLFALCFSCLLAHASNLVKLVRGCVSFSWT TILCIAIGCSLLQIIIATEYYTLIMIRGMMFVN MTPCQLNVDFVVLLYVLFLMALTFFVSKAT FCGPCENWKOHGRLIFIVVLFSIIIWVVMSML LRGNPCFGRQPQWDDPVVCIALVTNAWVFL LLYIVPELCILYRSCRGECPLQGNACPVTAYQ HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQVENQELSRDFWGCHLORDFGQSGAPG GRGAFRQEGFGSSWCKVLNGSGAPG HOUSE LLYIVPELCLLYRDGGGGAPU HSFQVENQELSRDKWKVLLNSDFLSHSGA HSFQLUNDERSGAPGGGGAPU LASKGELLRSFRSSELLLREGLQSSKLGK HSKLIAFFGLEGLGLACK LASKGELLRSFRSSELLLREGLLSSKLGK HSKLIAFFGLEGLGLACK LASKGELLRSFRSSECELLREGELLSK HSKLIAFFGLEGLGSKLGK HSKLIAFFGLEGLGLACK HSKLIAFFGLEGLGSKLGK HSKLIAFFGLEGLGSKLGK HSKLIAFFGLEGLGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK HSKLIAFFGCK	1294	2644	A	10129	91	1042	
GVLFALCESCLLAHASNI.VKI.VRGCVSFSWT TILCIAIGCSI.QIIIATEVYTLIMTRGMMFVN MTPCQLNVDFVVLLVYVLFMALTFFVSKAT FCGPCENWKQHGRLIFITVLFSIIWVVWISML LRGNPPGFQRQPWDDDFVVCLAVTNAWVFL LLYIVPELCILYRSCRQECPLQGNACPVTAYQ HSFQVENDELSRDK.WLINSDFLSHSGA 1295 2645 A 10133 376 518 RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG GRGAPRQEGPGSSWRQV 1296 2646 A 10135 3 551 EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE NIQDVLSALPNPDDYFLLRWLQARSFDLQKS EDMLRKHMEFRKQDAILLAWQARSFDLQKS EDMLRKHMEFRKQDAILLAWQARSFDLQKS EDMLRKHMEFRKQDAILLAWCPGELLQE 1297 2647 A 10138 48 407 MVSSCCGSVCSDQGCQDLQETCCRPSCCE 1298 2648 A 10156 94 453 GSNRKSAEMFSQVPTASGCYYLNSMTPEG QEMYLRFDQTTRSPYRMSRLARHQLVTKI QEMYLRFDQTTRSPYRMSRLARHQLVTKI QEMYLRFDQTTRSPYRMSRLARHQLVTKI QEMYLRFDQTTRSPYRMSRLARHQLVTKI QEMERKACDWLRARFQPYAQLVEDSOFP NIVAVKNDHDFLEKDLGEFLCRRLNT 1299 2649 A 10161 1 393 PRFSELVDGRGVSARFGGSPSKAATVRSQPT ASQLENMBEAPKRYSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPG KGRCCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VITYTGILGPGLIGNICALWATH VITYTGILGPGLIGNI	1	}	l		}		1
TILCIAIGCSILQIIIATEYYTLIMTRGMMFVN MTPCQLNVDFVVLLVYVLFMALTFFVSKAT FCGPCENWKQHGRLIFTVLFSIIIWVWISML LRGNPQFQRQWDDFVVCLAVTNAWVFL LLYIVVPELCLIYRSCRQECPLQGNACPVTAYQ HSFQVENQELSRDKWKVLLNSDFLSHSGA 1295 2645 A 10133 376 518 RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG GRGAPRQEGPGSSWRQV GRGAPRQEGPGSSWRQV 1296 2646 A 10135 3 551 EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE MIQDVLSALPPPDDFLRWLQARSFDLQKS EDMLRKHMEFRKQQDLANIL AWQPPEVVRL YNANGICCHDGEGSPVWYHIVGSQDPKGLL SASKQELLRDSFRSCELLLRECELQSQKLGKR VEKIIAIFGLEGLGJRDLWKPIGELLQE SASKQELLRDSFRSCELLLRECELQSQKLGKR VEKIIAIFGLEGLGJRDLWKPIGELLQE TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC GRYSCOPTCCRPSCCE TTCCRTTCCRTCCRTCCRTSCCVSSCCRPQCC TTCCRTTCCRTSCCVSSCCRPQCC GRYSCOPTCRPSCCE TTCCRTTCCRTSCCVSSCCRPQCC TTCCRTTCCRTSCCTTCCRTSCCTSCCRPQCC TTCCRTTCCRTSCCTSCRPQCC TTCCRTTCCRTSCCTSCCTTCCTTCTTSCTCRTSCCTSCRPQCC TTCCRTTCCRTSCCTSCCRPQCC TTCCRTTCCRTSCCTSCCTSCCRPQCC TTCCRTTCCRTSCCTSCCRPQCC TTCCRTTCCRTSCCTSCCTSCCRPQCC TTCCRTTCCRTSCCTSCCTSCCRPQCC TTCCRTTCCRTSCTTCTTSCTTCTTTCTTTTCTTTTCTT	1	1	l		}		
MTPCQLNVDFVULLYYVILMALTFFVSKAT FCGPCENWKQHGRLIFTVLFSIIIWVVWISML LRGNPQFQRQPQWDDPVVCIALVTNAWVFL LLYIVPELCILYRSCRQECFLQGNACFVTAYQ HSFQVENQELSRDKWKVLINSDFJSHSGA 1295 2645 A 10133 376 518 RPRVVTHINSQWCFLPQDHPGWLPGQSGAPG GRGAPRQEGFGSSWRQV 1296 2646 A 10135 3 551 EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE NIQDVLSALPNPDDYFLLRWLQARSPDLQKS EDMLRKHMEFRKQQDLANILAWQPPEVVRL YNANGICGHDGEGSFVWYHIVGSQDPKGLLL SASKQELLRDSFRSCELLLRECELQSQKLGKR VEKILAFFGLEGLGDLWKPGIELQE 1297 2647 A 10138 48 407 MVSSCCGSVCSDQGCGQDLCQETCCRPSCCE TTCCRTTCCRPSCCVSSCCRPQCCQ QVCCQPTCCRPSCCETTCCHPXCC QPVCCQFTCCRPSCCETTCCHPXCC QPVCCQFTCCRPSCCETTCCHPXCC QPVCCQFTCCRPSCCETTCCHPXCC QEWYLRFDQTTRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP NIVAVKNDHDFLEKDGEFLCRRINT 1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQFT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGTCVPGADAHSCDCGPGF KGRCELACKVSRPCTRLFSETKAFPVWEGG VCHIV VTYTGLGPGLIGNILALWYYGYMKETKRA VTYTGLGPCGRRAGSSISDLQGKGFKGTC EKHYPGYGSARHSPQASAGGSPWQGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAR LLCSGANRCKUVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPPLPDWVAAKTLAKA LRAPGRPWILAAPLDGLGGFGFSTAP RTLSVEEPGVECNQLCLYAPPLPDWVAAKTLAKA LRAPGRPWILAAPLDGLGGFGFSTAP RTLSVEEPGVECNQLCLYAPPLPDWVAAKTLAKA ARABSESLHVVDENKNESKELKHVHAKSEP	1	i		İ	İ	Į.	
FCGPCBNWKQHGRLEFITVLFSIIWVVWISML LRGNPQFQRQPQWDDPVVCIALVTNAWVFL LLYIVPELCILYRSCRQECPLQGNACPVTAYQ HSFQVENQELSRDKWKVLLNSDFLSHSGA 1295 2645 A 10133 376 518 RPRVVTHNSQWCFLPQDHPGWLPQSGAPG GRGAPRQEGPGSSWRQV 1296 2646 A 10135 3 551 EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE NQDVLSALPNPDDYFLLRWLQARSPDLQKS EDMLRKHMEFRKQQDLANILAWQPPEVVRL YNANGICGHDGGSPVWYHIVGSQDRGLLL SASKQBLLRDSFRSCELLLRECELQSQKLIGKR VEKHAIFGLEGLGLRDLWKPGELLQE VEKHAIFGLEGLGLRDLWKPGELLQE 1297 2647 A 10138 48 407 MYSSCCGSVCSQDGCGQDLCQETCCRPSCCE TTCCRTTCCPSSCCVSSCCRPQCCQSVCCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCTSCCVSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCVSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCTSCCTTCPSCCVSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCTSCCTTCCPSCCVSCCRPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCTSCCTTCCPSCCTSCCTPQCCQSVCQPT CSRPSCQTTCCRTTCCPSCCTSCCTTCCTSCCTSCCTPQCCQSVCQPT CSRPSCQTTCCTTTCCPSCCTTCPSCCTSCCTTCCTTCCPSCCTTCTTCTTCPSCTTCPSCTTCPSCTTCTTTCT	1	ļ	.		j		,
LRGNPQFQRQWDDPVVCIALVTNAWVFL		ł	l '	İ			
LLYIVPELCILYRSCRQECPLQGNACPYTAYQ			1]		
HSRQVENQELSRDKWKVJLNSDFLSHSGA	1	1					
1295	ĺ	1		1 .	[LLYIVPELCILYRSCRQECPLQGNACPVTAYQ
1296 2646 A 10135 3 551 EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE MIQDVLSALPPDDYFLLRWLQARSFDLQKS EDMLRKHMEFRKQDDLANILAWQPPEVVRL YNANGICGIBGEGSPVWYHIVGSQDFKGLLL SASKQELLRBSFRSCELLIRECELQSQKLGKR VEKILAIFGLEGLQCORDLOGETCCRPSCCE TTCCRTTCCRPSCCVSSCCRPQCCQVCCQPT CSRPSCQTTCCRTTCYRPSCVSSCCRPQCCQVCCQPT CSRPSCQTTCCRTTCYRPSCVSSCCRPQCCQVCCQPT CSRPSCQTTCCRTTCYRPSCVSSCCRPQCCQ QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP NIVAVKNDHDFLEKDLGEPLCRRLNT 1299 2649 A 10161 1 393 PRFSELVDGRGVSARFGGSPSKAATVRSQPT ASAQLENMEBAPKRVSLALQLPEHGSKDIGN VPGNCSSNPCQGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV VPGNCSSNPCQGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV							HSFQVENQELSRDKWKVLLNSDFLSHSGA
1296	1295	2645	Α	10133	376	518	RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG
NIQDVLSALPNPDDYFLLRWLQARSFDLQKS EDMLRKHMEFRKQDLANILAWQPPEVVRL YNANGICHDGEGSPVWFIUVSQDPRGILLL SASKQELLRDSFRSCELLLRECELQSQKLGKR YEKILAFGLEGLGLRDLWKPGIELLQE			}				GRGAPRQEGPGSSWRQV
EDMLRKHMEFRKQQDLANILAWQPPEVVRL YNANGICGHDGGSPVWYHVGSQDPKGLLL SASKQELLRDSFRSCELLLREGLQSQKLGKR VEKILAIFGLEGLQSVLLGKE VEKILAIFGLEGLQLRDLWKPGIELLQE TTCRTTCCRTSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVCQPTCRPSCCTTCHPXCC QPVCQPTCRPSCCTTCHPXCC QPVCQPTCRPSCQPTCHPXCPTCHPXCTTTCHPACTCHPXCC QPVCQPTCRPSCQPTCHP	1296	2646	Α	10135	3	551	EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE
YNANGICGHDGEGSPVWYHIVGSQDPKGLLL SASKQELLRDSFRSCELLLRECELQSQKLGKR VEKILAIFGLEGLGLADL WKPGIELLQE	1	ļ	1	l			NIQDVLSALPNPDDYFLLRWLQARSFDLQKS
SASKQELIRDSFRSCELLIRECELQSQKLGKR			1				EDMLRKHMEFRKQQDLANILAWQPPEVVRL
VEKITÄIFGLEGLGLRDLWKPGIELLQE		•	1				YNANGICGHDGEGSPVWYHIVGSQDPKGLLL
1297 2647 A 10138 48 407 MVSSCCGSVCSDQGCQDLCQETCCRPSCCE TTCCRTTCCRTSCCVSSCCRPQCCQSVCCQPT CSRPSCCVTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRTSCCYSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC 1298 2648 A 10156 94 453 GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT 1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVOSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		ŀ	·		į		SASKQELLRDSFRSCELLLRECELQSQKLGKR
TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRTTCCTPXCC 1298 2648 A 10156 94 453 GGRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT ASAQLENMEEAPKRYSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVOSARHSPQASAGGSPWORGKAQT RWLGKPDGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSSLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV			l	1	ļ		VEKIIAIFGLEGLGLRDLWKPGIELLQE
TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRTCCHPXCC 1298 2648 A 10156 94 453 GGNKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHYPGVGSARHSPQASAGGSPWORGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1297	2647	A	10138	48	407	MVSSCCGSVCSDQGCGQDLCQETCCRPSCCE
QPVCCQPTCCRPSCCETTCCHPXCC			1				
QPVCCQPTCCRPSCCETTCCHPXCC	1						CSRPSCCOTTCCRTTCYRPSCCVSSCCRPQCC
1298 2648 A 10156 94 453 GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRINT 1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV	1	ł	1		ł	Ì	
QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRINT 1299 2649 A 10161 I 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1298	2648	A	10156	94	453	
QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT 1299 2649 A 10161 I 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV			1	1	1		1
INIVAVKNDHDFLEKDLGEPLCRRLNT 1299 2649 A 10161 I 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT		1					((
1299 2649 A 10161 I 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		1					
ASAQLENMEEAPKRVSLALQLPEHGSKDIĞN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1299	2649	A	10161	1	393	
VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1		١	10101	1 -	-,,,	1
KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		1	l			l	
1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV]]]		}		-
1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV			1				·
VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1300	2650	<u> </u>	10162	08	201	
VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1300	2030	^	10102	100	1 371	
1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV							i l
1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV]]	}]	1]	
KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1201	2651	 	10165	1	7545	1 <u></u>
EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1301	2031	A	10102) ¹	/343	
RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1	1	1	1			
LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV			1		}	1	
PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		j	ļ	j			
LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1	ļ	1				
RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		ĺ					•
DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1		1		1		
ARRLSESLHVVDENKNESKIEREHKRRTSTPV	Ì	1	1		1	1	
	İ	1	1				
		L		<u></u>	L	L	IMEGVQEETDTRDVKRQVERSEICTEEPQKQ

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
denos			717	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i			ľ	peptide	Sequence	/=possible nucleotide deletion, \=possible
						nucleotide insertion
<u> </u>				sequence	ļ	
						KSTLKNEKHLKKDDSETPHLKSLLKKEVKSS
						KEKPEREKTPSEDKLSVKHKYKGDCMHKTG
			ĺ	i		DETELHSSEKGLKVEENIQKQSQQTKLSSDDK
, 1						TERKSKHRNERKLSVLGKDGKPVSEYIIKTDE
						NVRKENNKKERRLSAEKTKAEHKSRRSSDSK
						IQKDSLGSKQHGITLQRRSESYSEDKCDMDST
			l			NMDSNLKPEEVVHKEKRRTKSLLEEKLVLKS
. 1			1			KSKTQGKQVKVVETELQEGATKQATTPKPD
. 1			[KEKNTEENDSEKORKSKVEDKPFEETGVEPV
. 1						LETASSSAHSTQKDSSHRAKLPLAKEKYKSD
						KDSTSTRLERKLSDGHKSRSLKHSSKDIKKKD
. !						ENKSDDKDGKEVDSSHEKARGNSSLMEKKL
,	·]	ļ		SRRLCENRRGSLSQEMAKGEEKLAANTLSTP
, I						SGSSLQRPKKSGDMTLIPEQEPMEIDSEPGVE
,						
					1	NVFEVSKTQDNRNNNSHQDIDSENMKQKTS
						ATVQKDELRTCTADSKATAPAYKPGRGTGV
						NSNSEKHADHRSTLTKKMHIQSAVSKMNPGE
, i			1			KEPIHRGTTEVNIDSETVHRMLLSAPSENDRV
. 1	i '					QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS
						LSSVTVVPLRESYDPDVIPLFDKRTVLEGSTA
			ĺ			STSPADHSALPNQSLTVRESEVLKTSDSKEGG
			}			EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ
. 1						GKVIMPLGSKLTGVIVENENITKEGGLVDMA
						KKENDLNAEPNLKQTIKATVENGKKDGIAVD
						HVVGLNTEKYAETVKLKHKRSPGKVKDISID
						VERRNENSEVDTSAGSGSAPSVLHQRNGQTE
. 1						DVATGPRRAEKTSVATSTEGKDKDVTLSPVK
						AGPATTTSSETRQSEVALPCTSIEADEGLIIGT
						HSRNNPLHVGAEASECTVFAAAEEGGAVVTE
, J					i	GFAESETFLTSTKEGESGECAVAESEDRAADL
						LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE
						KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG
						TEIRAGSISSEEVDGSQGNMMRMGPKKETEG
						TVTCTGAEGRSDNFVICSVTGAGPREERMVT
						GAGVVLGDNDAPPGTSASQEGDGSVNDGTE
, [GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS
	5			. 1		SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH
, 1				,		
. 1						ASTCTGLGEESEGVLICESAEGDSQIGTVVEH
				. 1		VEAEAGAAIMNANENNVDSMSGTEKGSKDT
				`		DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE
J				ļ		GPMTSAASDQSDSQLEKVEDTTISTGLVGGS
				İ		YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE
				ŀ		NEECDGLMATTASGDITNQNSLAGGKNQGK
Í				ļ		VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE
ļ		į		l		ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS
j						RSEEKDECAMISTSIGEEFELPISSATTIKCAES
				ļ		LQPVAAAVEERATGPVLISTADFEGPMPSAPP
				l		EAESPLASTSKEEKDECALISTSIAEECEASVS
				l	1	GVVVESENERAGTVMEEKDGSGIISTSSVEDC
				ļ	•	EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS
}				1	İ	TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA
				l	l	ALISTSTAECMPISASIDRHEENQLTADNPEGN
				ļ		GDLSATEVSKHKVPMPSLIAENNCRCPGPVR
1					1	GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH
				l	Į.	DOVELOL AFWASTEEGHINGLAAHKESAGOOH
				ļ	,	PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL
8						HLINAEEKNVLLNSLQKEDKSPETGTAGGSST
				1		ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK
			. 1		Į.	DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA
j		1	ı j	J		EHSFLPAEQQGSEDNLKTSTTKCITGQESKIAP

NO. of NO. of nucleic enticle sequence NO. of No. of N	050 10	OF O ID	34-4	000	Deadistad	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
mucleotide seq	SEQ ID	SEQ ID	Met	SEQ	Predicted		Amino acid Sequence (A-Alainie C-Cysteine,
Sequence		-	nou				F-Phenylelenine G-Glycine H-Histidine
1302 2652 A 10167 321 842 EPSLEPTER ENTSET INSTITION							
1914 ng to first amino acid de residue of peptide		•			*		1=150/eucine, K=Lysine, L-Leucine,
amino acid resided of peptide resided of peptide resided of peptide sequence T-Thronine, V-Valine, W-Tryptophan, Y-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide sequence T-Thronine, V-Valine, W-Stop codon, peptide, V-Thronine, V-Valine, W-Stop codon, peptide, V-Thronine, V-Valine, W-Stop codon, peptide, V-Thronine, V-Valine, W-Thronine,		uence					M=Methonine, N=Asparagine, r=roine,
Pepside of peptide sequence	uence		[· {	914			T-Thursday V-Voline W-Trustophen
peptide							1=Inreonine, v=valine, w=Itypiopilaii,
			ł i			sequence	Y=1yrosine, X=Unknown,stop codon,
SITIMIPPATYSVALLARKCEQDLTIKNDYSOK WTOQASAEKTQDINSTKSSPEEGOIMVIVS SEENVCDIONESPE.NVLGGLK.KANLKMEA VPTSEEEKNGEIL.APPSELOGGKPSGLABLQRE PLLVNESI.NVENSGFRTNEERISSSYNKGBISS GRKDNARAISGHSVADPKEVEEEREHMÄKR KRKQHYI.SSEDEPODNPDVLDSRIETAQRAC PTEPHATTEENSRIDE.EILKYSSSTRISTISTS MEEKDEYSSSETTGEKPEONDDTIKSQE PTEPHATTEENSRIDE.EILKYSSSTRISTISTS MEEKDEYSSSETTGEKPEONDDTIKSQE PTEPHATTEENSRIDE.EILKYSSSTRISTISTS MEEKDEYSSSETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDDTIKSQE PSESETTGEKPEONDSPEPLEPLLADELE DGFGGHFFYHCLVAEVREMENSFEPLEPLLADELE DGFGGHFFYHCLVAEVREMENSFEPLEPLLADELE DGFGGHFFYHCLVAEVREMENSFEPLEPLLADELE DGFGGHFFYHCLVAEVREMENSFEPLEPLLADELE DGFGGHFFYHCLVAEVREMENSFEPLEPLLADELE DGFGGHFFYHCLVAEVREMENSFEPLEPLADELE DGFGGHFFYHCLVAEVREMENSFEPLEPLADELE DGFGGHFFYHCLVAEVREMENSFEPLEPLADELE DGFGGHFFYHCLVAEVREMENSFEPLEPLADELE DGFGGHFFYHCLVAEVREMENSFEPLE DGFGGHFFYHCLVAEVREMENSFEPLE DGFGGHFFYHCLVAEVREMENSFEPLE DGFGGHFFYHCLVAEVREMENSFEPLE DGFGGGGFFFYHCLVAEVREMENSFEPLE DGFGGGGFFFTHADELE DGFGGGGGFFFTHADELE DGFGGGGGFFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGFFTHADELE DGFGGGGGGFFTHADELE DGFGGGGGGFTHADELE DGFGGGGGGFFTHADELE DGFGGGGGGFTHADELE DGFGGGGGGFTHADELE DGFGGGGGGFTHADELE DGFGGGGGGFTHADELE DGFGGGGGGFTHADELE DGFGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG							
WTTQASAEKTGDDNSTRKSFPEGDIMVTVS SEENVCLOINGESFILAN/LGGLIKANALIKMEA					sequence		
							SHIMIPPALISVALLAFACEQULITANDISUA
1302 2652 A 10167 321 842 EPSLEPFILESPRAYERS SERVEN SERVE							
PILLVNESLNVENSGRTNEEHISESYNKOES RRKDNAEAISGHVADPKEVEEERERHMPKR	•		[[ĺ		
	i '		i 1				
1302 2652 A 10167 321 842 EPSLEPLATKERNSDLEELPKTSENSTTSRV MEEKDEYSSETTIGEKPRONDDDTIKSOG PSTLEPHATKERNSDLEELPKTSENSTTSRV MEEKDEYSSSETTIGEKPRONDDDTIKSOG PSTLEPPLREPPREPPREPPEPEPEPELAGPEPEL AGPEPH PVFYFFLSVVHPPKELAKYEYMEEQVILTEKG NSTVAGRGTSVRCLSFSRRLPFLLPLLALDLE DGGGHPFYHCLVAEVKERWTYEGAPSYFP EARETKCYVRSSVGCVEPLTIQAEVTENLDR KNSQVFKLLKKK NSTVAGRGTSVRCLSFSRRLPFLLPLLALDLE DGGGHPFYHCLVAEVKERWTYEGAPSYFP EARETKCYVRSSVGCVEPLTIQAEVTENLDR KNSQVFKLLKKK NSTVAGRGTSVRCLSFSRRLPFLLPLALDLEURKGA KNSQVFKLLKKK NSTVAGRGTSVRCLSFSRRLPFLLLDSVRKGA KNSQVFKLLKKK NSTVAGRGTSVRCLSFSFRPLPFLLDSVRKGA KNSQVFKLLKKK NSTVAGRGTSVRCLSFSFRPLPFLLDSVRKGA RACEICITYI LCTLSPGISGTAGSCLTTEPGTELGTSFAQNGF YHEAVVLFTQALKINPOPHRLFGRRSFCHE LGQPAWALADAQVALTLRGGWSGFGAPGALQPLPHA ELAPSGLPSIR.CPRSTALRSPGI.SPLLH ELAPSGLPSIR.CPRSTALRSPGI.SPLLT >f		1					
PETEPHATKEENSRÜLEELPKTSSEINSTTSRV	·						= -
MERKDEYSSSTTIGER/PONDDITIKSQE	ļ,		}		j		
1302 2652 A 10167 321 842 EPSLPFILRSPARPPRPAFFSPELAGPEPT FVFYFILSYVERBEALXYPERGVILTEKG NSTVAGRGTSVRCLSPSPRPLPFLLADLLE DGRGEHPFYHCLVAEVPKEHWTPEGNPSFP EARETKCYVRSSVGCVEPLITQAEVTENLDR KNSQVYFLLKKK NSQVYFLLKKK NSQVYFLLKKK NSQVYFLLKKK NSQVYFLLKKK NSQVYFLLKKK NSQVYFLLKKK NSQVFKLLKK NSQVFKLLKK NSQVFKLKK NSQVFKLKK NSQVFKLKK NSQVFKLKK NSQVFKLKK NSQVFKLKK NSQVFKLKK NSQVFLKK NSQVFLKK NSQVFLKK NSQVFLKK NSQVFLK NSQVF							
PFYFYFISYMPPKELAKYEYMEEQVILTEKG NSTVAGRGTSYRCLSPSPRPLIPLIALDLE DGFGEHPFYHCLVAEVPKEHWTPEGNPSPFP EARETKCTYRSSVGCVSPLTTQAEVTENLDR KNSQQVFKLLKKK NSMLLKKRELLINSLGEGTINGLLDELLETNV LSQEDTETVKCENVTUDKARDLLDSVIRKGA RACEICITY	1200	0650		10167	201	942	
NSTVAGRGTSVRCLSFSRRIPPLILADILIE DOPGCHPFYPICLVAEVPKEHWTPEGNPSPPP	1302	2652	A	10167	321	842	
DGFGEHFYHGLVAEVPKEHWITEGNPSPP	1						
EARSTKCYVRSSVGCVEPLTTQAEVTENLDR	1				•		
1303 2653 A 10171 206 429 NMILKKRRLINSLGEGTINGLLDELLETINV LSQEDTEIVKCENVTVIDKARDLLDSVIRKGA RACEICITY	ļ		[]				
1303 2653 A 10171 206 429							
LSQEDTEIVKCENTVIDKARDLLDSVIRKGA RACEICITY						400	NNSQQVFKLLKKK
RAČEICITY 1304	1303	2653	A	10171	206	429	
1304		ł					
YHEAVVLFTQALKLNPQDHRLFGNRSFCHER LGQPAWALADAQVALTLRPGWPRGLFLGK ALMGLQRFREAAAVFQETLRGGSQPDAAREL RSCLLHLTLQGQRGGICAPPLSPGALQPLPHA ELAPSGLPSLCRFALRSFGLSPLLH TIOLGRRFRAVDGAAMAACEGRRSGALGSSQ SDFLTPPVGGAPWAVATTVMYPPPPPBUR DFISVTLSFGESYDNSKSWRRSVKRKWKWQL SRLQRNMILFLLAFLLFCGLLFYINLADHWKG RNTCT						1.501	
LGQPAWALADAQVALTLRPGWPRGLFRLGK ALMGLQRFREAAAVQETLRGGSQPDAAREL RSCLLHLTLQQRGGICAPPLSPGALQPLPHA ELAPSGLPSLRCPRSTALRSPGLSPLLH	1304	2654	A	10184	970	1524	
ALMGLQRREAAAVFOETLRGGSQPDAAREL RSCLLHLTLQQRGGICAPPLSPGALQPLPHA ELAPSGLPSLRCPRSTALRSPGLSPLLH 1305 2655 A 10194 2 394 TDLLGRRFRVDGAAMAACEGRRSGALGSSQ SDFLTPPVGGAPWAVATTVVMYPPPPPPPHR DFISVTLSFGESYDNSKSWRRRSCWRKWKQL SRLQRINMLFLLAFLLFCGLLFYINLADHWKG RINTCT 1306 2656 A 10195 1 410 IPGSTISLEGPLSKWTNVMKGWQYRWFVLDY NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI GIDDEDSSTFITIVDQKTFHFQARDADEREK WIHALEETILRHILQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK WIHALEETILRHILQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK WIHALEETILRHILQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK SPDQQNYTKSR SPDQQNYTKSR SPDQQNYTKSR SPDQQNYTKSR SPDQQNYTKSR SPDQQNYTKSR SPDQNYTKSR SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQWISLHTQYLESFLRSQFYMLRMDGPL PLPYRHVIAIMAAARHQCSYLINM RGWPEQQSTGRPEDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKILCGGSILHWWVVLTAAHCFRRTLLDM AV LPGADYGGHLSLRLFHLLLTSAAWVPDESQ VILNSALCVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKQGIPDVVVGDH DRDCDSHPGKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRFTI KAG STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	i		1 '		1	[
RSCLLHLTLQQQRGGICAPPLSPGALQPLPHA	ļ					İ	LGQPAWALADAQVALTLRPGWPRGLFRLGR
ELAPSGLPSLRCPRSTALRSPGLSPLLH 1305	1	1	1		1		
1305		•			1		
SDFLTPPVGGAPWAVATTVVMYPPPPPPPR	L	<u> </u>					
DFISVTLSFGESYDNSKSWRRRSCWRKWKQL SRLQRNMILFILAFLLFCGLLFYINLADHWKG IRNTCT	1305	2655	A	10194	2	394	
SRLQRNMILFLLAFLLFCGLLFYINLADHWKG IRNTCT			j		}		
IRNTCT	j	ļ			1		
1306	ļ	1			1		,
NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI GIDDEDDSTFTITVDQKITFHFQARDADEREK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK 1307 2657					ļ		
GIDDEDDSTFTITVDQKTFHFQARDADEREK	1306	2656	A	10195	[]	410	· ·
WIHALEETILRHTLQLQVRVFTWFPDSSLVGA		1					
FFFWLVSGFFFK	ł	l	{			1	GIDDEDDSTFITTVDQKTFHFQARDADEREK
1307 2657	1		·		·		
DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS SPDQQNYTKSR			<u> </u>				
SPDQQNYTKSR	1307	2657	A	10205	85	308	QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM
1308 2658 A 10214 2 453 ECGGIRQPGPGPPPALASAPAATMINRVGGSPS AAANYLLCTINCRK VLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDERTINFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	i						
AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	Ĺ	<u> </u>					
SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRFTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1308	2658	A	10214	2	453	
DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1	1	1	ł	ŀ		
PLPYRHYIAIMAAARHQCSYLINM 1309 2659 A 10233 45 421 RGWPEQQSTGPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP			1				
1309 2659 A 10233 45 421 RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP			}]	1		
RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP			<u> </u>			ļ	
RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1309	2659	A	10233	45	421	RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP
SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP]	1	}	j	1		
AV 1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRKK HKKGQSAEIQKKRTRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP				ļ			
1310 2660 A 10241 243 442 AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	l .	1	1				
HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	L		<u> </u>			<u> </u>	
MAK 1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1310	2660	A	10241	243	442	
1311 2661 A 10261 751 176 LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1	}		l		1	
VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1	1	1				1
VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1311	2661	A	10261	751	176	
QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP							
QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP			l	1			LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF
DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP]		1	l	1	1	
QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG 1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP		ļ	ļ		1		
KAG KAG	1	}	1	Į.	1	1	
1312 2662 A 10270 3 669 STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP	1			1	1	}	
	1312	2662	A	10270	3	669	
	1		1])	***	SMTILDKKDGEQAKALFEKVRKFRAHVEDSD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LIYKLYVVQTVIKTAKFIFILCYTANFVNAISF EHVCKPKVEHLIGYEVFECTHNMAYMLKKL
1313	2663	Ā	10287	1221	266	LISYISIICVYGFICLYTLFWLFRIPLKEYSFEKV REESSFSDIPDVKNDFAFLLHMVDQYDQLYS KRFGVFLSEVSENKLREISLNHEWTFEKL
						GAHRVLSPAQGAQPRLRSAASVEVSMVGQR VLLLVAFLLSGVLLSEAAKILTISTLGGSHYLL LDRVSQILQEHGHNVTMLHQSGKFLIPDIKEE EKSYQVIRWFSPEDHQKRIKKHFDSYIETALD GRKESEALVKLMEIFGTQCSYLLSRKDIMDSL KNENYDLVFVEAFDFCSFLIAEKLVKPFVAIL PTTFGSLDFGLPSPLSYVPVFPSLLTDHMDFW GRVKNFLMFFSFSRSQWDMQSTFDNTIKEHF PEGSRPVLSHLLLKAELWFVNSDCAFDFARPL LPNTVYIGGLMEKPIKPVPQVSEPSAFSLGFT
1314	2664	A	10288	536	1890	NVQLAKFSSTLVFFFSCDADPSALAKYVLAL VKKDKSEKELKALCIDQLDVFLQKETQIFVEK LFDAVNTKSYLPPPEQPSSGSLKVEFFPPQEK DIKKEEITKEEEREKKFSRRLNHSPPQSSSRYR ENRSRDERKKDDRSRKRDYDRNPPRRDSYRD RYNRRGRSRSYSRSRSRSWSKERLRERDRD RSRTRSRSTRSRERDLVKPKYDLDRTDPLEN NYTPVSSVPSISSGHYPVPTLSSTITVIAPTHHG NNTTESWSEFHEDQVDHNSYVRPPMPKKRC RDYDEKGFCMRGDMCPFDHGSDPVVVEDVN LPGMQPFPAQPPVVEGPPPPGLPPPPILTPPPV NLRPPVPPPGPLPPSLPPVTGPPPPLPPSG MDAPPNSATSSVPTVVTTGIHHQPPPAPPSLFT ADTYDTDGYNPEAPSITNTSRPMYRHRVHPR AKLG
1315	2665	A	10293	447	1331	SHPLLSCPEKVSAKLRAAAEAAAEERRTRGA GSRGICAGLRSVAPGPEPLKQEEGRREWGSSI GTPSPCGSAQAAAAAAAEEATEKIPALRPALL WALLALWLCCATPAHALQCRDGYEPCVNEG MCVTYHNGTGYCKCPEGFLGEYCQHRDPCE KNRCQNGGTCVAQAMLGKATCRCASGFTGE DCQYSTSHPCFVSRPCLNGGTCHMLSRDTYE CTCQVGFTGRNPKCPGGNLNYQFNGIIVVYS GGSVPPSGTKTSKPAEHNAMGTGSKNFASGT LWVMVSGATSTSTSTL
1316	2666	Α	10294	118	572	SLSMESNHKSGDGLSGTQKEAALRALVQRTG YSLVQENGQRKYGGPPPGWDAAPPERGCEIFI GKLPRDLFEDELIPLCEKIGKIYEMRMMMDF NGNNRGYAFVTFSNKVEAKNAIKQLNNYEIR NGRLLGVCASVDNCRLFVGGIPKTKK
1317	2667	A	10301	158	1956	LLKSCGVLLSGVCIPCEGKGPTVLVIQTAVPQ DRPTKSSMRSAAKPWNPAIRAGGHGPDRVRP LPAASSGMKSSKSSTSLAFESRLSRLKRASSE DTLNKPGSTAASGVVRLKKTATAGAISELTES RLRSGTGAFTTTKRTGIPAPREFSVTVSRERSV PRGPSNPRKSVSSPTSSNTPIPTKHLRTPSTKP KQENEGGEKAALESQVRELLAEAKAKDSEIN RLRSELKKYKEKRTLNAEGTDALGPNVDGTS VSPGDTEPMIRALEEKNKNFQKELSDLEEENR VLKEKLIYLEHSPNSEGAASHTGDSSCPTSITQ ESSFGSPTGNQLSSDIDEYKKNIHGNALRTSG SSSSDVTKASLSPDASDFEHITAETPSRPLSSTS NPFKSSKCSTAGSSPNSVSELSLASLTEKIQKM EENHHSTAEELQATLQELSDQQQMVQELTAE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutarnic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutarnine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion NEKLVDEKTILETSFHQHRERAEQLSQENEKL MNLLQERVKNEEPTTQEGKHELEQKCTGILE QGRFEREKLLNIQQQLTCSLRKVEEENQGAL EMIKRLKEENEKLNEFLELERHNNNMMAKTL EECRVTLEGLKMENGSLKSHLQG GECFIMAAVVQQNDLVFEFASNVMEDERQL
						GDPAIFPAVIVEHVPGADILNSYAGLACVEEP NDMITESSLDVAEEEIIDDDDDDITLTVEASCH DGDETIETIEAAEALLNMDSPGPMLDEKRINN NIFSSPEDDMVVAPVTHVSVTLDGIPEVMETQ QVQEKYADSPGASSPEQPKRKKK
1319	2669	A	10322	169	654	MEVRMSGSVAVTRAIAVPGLLLLIIATALSL LIGAKSLPASVVLEAFSGTCQSADCTIVLDAR LPRTLAGLLAGGALGLAGALMQTLTRNPLAD PGLLGVNAGASFAIVLGAALFGYSSAQEQLA MAFAGALVASLIVAFTGSQGGGQLSPVRLTL AGVXL
1320	2670	A	10323	441	2	KMNQVAVVIGGGQTLGAFLCHGLAAEGYRV AVVDIQSDKAANVAQEINAEYGESMAYGFG ADATSEQSVLALSRGVDEIFGRVDLLVYSAGI AKAAFISDFQLGDFDRSLQVNLVGYFLCARE FSRLMIRDGIQGRIIQINSKSDE
1321	2671	A	10332	1	453	RHRTAGPGSTISSRTDSASAPAARAMPCEYTY AKLTSDCSRPSLQWYTRAQSKMRRPRLLLKD ILKCTLLVFGVRILYILKLNYTTEECDMKNMH YVDPDHVKRAQKYAQQVLQKESPPKFAKTS MALLFEHRYSVDLLPFVQKAPTDSEA
1322	2672	A	10333	25	423	EPSNGPVVYSALGNEDDEILLLGKDIIGTFAAS ERKMRAHQVLTFLLLFVITSGASENASTSRGC GLDLLPQNVYLCDLDAIWGIVVEAVAGAGA LITLLMLILLGRLPFIKEKEKKSPAVLHFLFL LGTLG
1323	2673	A	10334	52	426	SSLGNEDDEILSLAKDITGMFVASHRKMRAH QVLTFLLLFVITSVASENASTSRGCGLDLLPQ YVSLCDLDAIWGIVVEAAGAGALITLLLMLI LLVRLPFFKEKEKKSPVGLHFLFLLGTLGP
1324		A,	10336	1	932	ERLCFPCMQSKIYSYMSPNKCSGMRFPLQEE NSVTHHEVKCQGKPLAGIYRKREEKRNAGN AVRSAMKSEEQKIKDARKGPLVPFPNQKSEA AEPPKTPPSSCDSTNAAIAKQALKKPIKGKQA PRKKAQGKTQQNRKLTDFYPVRRSSRKSKAE LQSEERKRIDELIESGKEEGMKIDLIDGKGRG VIATKQFSRGDFVVEYHGDLIEITDAKKREAL YAQDPSTGCYMYYFQYLSKTYCVDATRETN RLGRLINHSKCGNCQTKLHDIDGVPHLILIAS RDIAAGEELLYDGDRSAATAGSPGPRPRLATGWI
1325	2675	A	10338	3	870	PGSTISCSELKGTQCRATAGSRGRRPPMTCWL RGVTATFGRPAEWPGYLSHLCGRSAAMDLG PMRKSYRGDREAFEETHLTSLDPVKQFAAWF EEAVQCPDIGEANAMCLATCTRDGKPSARML LLKGFGKDGFRFFTNFESRKGKELDSNPFASL VFYWEPLNRQVRVEGPVKKLPEEEAECYFHS RPKSSQIGAVVSHQSSVIPDREYLRKKNEELE QLYQDQEVPKPKSWGGYVLYPQVMEFWQG QTNRLHDRIVFRRGLPTGDSPLGPMTHRGEE DWLYERLAP
1326	2676	A	10344	2	984	ARAAAHCGICRLVRWWRKRRSVMGIQTSPV LLASLGVGLVTLLGLAVGSYLVRRSRRPQVT LLDPNEKYLLRLLDKTTVSHNTKRFRFALPTA

SEQ ID SEQ ID Met SEQ Predicted Predicted end Amino acid sequence (A=Ala NO: of NO: of hod ID NO: beginning nucleotide D=Aspartic Acid, E=Glutami nucleotide peptide in nucleotide location F=Phenylalanine, G=Glycine, eotide seq- USSN location corresponding l=Isoleucine, K=Lysine, L=14	c Acid,
nucl- peptide in nucleotide location F=Phenylalanine, G=Glycine,	o raciu,
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
eotide seq- USSN location corresponding l=Isoleucine, K=Lysine, L=L	
seq- uence 09/496 correspondi to last amino M=Methionine, N=Asparagin	
uence 914 ng to first acid residue Q=Glutamine, R=Arginine, S	
amino acid of peptide T=Threonine, V=Valine, W=	Tryptophan
residue of sequence Y=Tyrosine, X=Unknown, *=	=Stop codop
peptide /=possible nucleotide deletion	\=nossible
sequence nucleotide insertion	, v possible
HHTLGLPVGKHIYLSTRID	GST VIRPVTPVTSD
EDQGYVDLVIKVYLKGVF	
LDSLKVGDVVEFRGPSGL	LTYTGKGHENIOP
NKKSPPEPRVAKKLGMLA	GGTGITPMI OLIRA
ILKVPEDPTQCFLLFANQT	EKDIII REDI EELO
ARYPNRFKLWFTLDHPPK	DWAYSKGFVTAD
MIREHLPAPGDDVLVLLC	GPPPMVOLACHPN
LDKLGYSQKMRFTY	
1327 2677 A 10345 1 968 LQSAGEGVTHVLILLESPA	RPVAAVTOVORR
RYHRLSDMSMLAERRRKO	
SNDDSKFGQRMLEKMGW	SKGKGLGADEOG
ATDHIKVQVKNNHLGLGA	TINNEDNWIAHO
DDFNQLLAELNTCHGQET	TDSSDKKEKKSES
LEEKSKISKNRVHYMKFTI	KGKDLSSRSKTDL
DCIFGKRQSKKTPEGDASF	
TIQEYFAKRMAALKNKPQ	
RKRGKKRNKEATGKDVE	
_ KPERAEAQERVAKKKSAP	
SSKASAQDAGDHVQPA	
1328 2678 A 10346 173 439 GSAAMKVKIKCWNGVAT	WLWVANDENCGI
CRMAFNGCCPDCKVPGDI	CPLVWGOCSHCE
HMHCILKWLHAQQVQQH	CPMCROEWKFKE
1329 2679 A 10351 3 964 QMEPGNDTQISEFLLLGFS	
SMYLVTVLGNLLIILATISI	DSHLHTPMYFFLSN
LSFADICVTSTTIPKMLMN	
MQMYFFILFAGFENFLLSV	
LHYMVIMNPHLCGLLVLA	SWTMSALYSLLOI
LMVVRLSFCTALEIPHFFC	ELNOVIOLACSDSF
LNHMVIYFTVALLGGGPL	
AISSAQGKYKAFSTCASHL	SVVSLFYGAILGV
YLSSAATRNSHSSATASVN	
YSLRNKDIKRALGIHLLWO	GTMKGQFFKKCP
1330 2680 A 10352 34 2573 IPFLKSCCCCCLFDFPPPPL	DQVQEEECEVERV
TEHGTPKPFRKFDSVAFGE	
TDPPNWQQLVSREVLLGL	KPCEIKRQEVINEL
FYTERAHVRTLKVLDQVF	YQRVSREGILSPSE
LRKIFSNLEDILQLHIGLNE	QMKAVRKRNETS
VIDQIGEDLLTWFSGPGEE	KLKHAAATFCSNQ
PFALEMIKSRQKKDSRFQT	FVQDAESNPLCRR
LQLKDIIPTQMQRLTKYPL	LLDNIATYTEWPT
EREKVKKAADHCRQILNY	VNQAVKEAENKQ
RLEDYQRRLDTSSLKLSEY	
RKMIHEGPLVWKVNRDKI	
LLQKQDDRLVLRCHSKILA	ASTADSKHTFSPVI
KLSTVLVRQVATDNKALF	VISMSDNGAQIYE
LVAQTVSEKTVWQDLICR	Maasvkeqstkpi
PLPQSTPGEGDNDEEDPSK	
QSPDRDLGLESTLISSKPQS	SHSLSTSGKSEVRD
LFVAERQFAKEQHTDGTLI	KEVGEDYQIAIPDS
HLPVSEERWALDALRNLG	LLKQLLVQQLGLT
EKSVQEDWQHFPRYRTAS	QGPQTDSVIQNSE
NIKAYHSGEGHMPFRTGTO	
SFAPRDSVGLAPQDSQASN	
MPTMEPEGGLDDSGEHFFI	DAREAHSDENPSE
GDGAVNKEEKDVNLRISG	NYLILDGYDPVQE
, , , , , , , , , , , , , , , , , , , ,	PAVESTHOOOHSP
SSTDEEVASSLTLQPMTGIE	111201112021101
SSTDEEVASSLTLQPMTGII QNTHSDGAISPFTPEFLVQC	QRWGAMEYSCFEI
SSTDEEVASSLTLQPMTGIE	QRWGAMEYSCFEI EADLEHLKKVEE

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end of nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1331	2681	A	10353	1	2100	AVEFAEGALTMAPWPELGDAQPNPDKYLEG AAGQQPTAPDKSKETNKTDNTEAPVTKIELLP SYSTATLIDEPTEVDDPWNLPTLQDSGIKWSE RDTKGKILCFFQGIGRLILLLGFLYFFVCSLDIL SSAFQLVGGKMAGQFFSNSSIMSNPLLGLVIG VLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIP IIMGANIGTSITNTIVALMQVGDRSEFRRAFA GATVHDFFNWLSVLVLLPVEVATHYLEIITQL IVESFHFKNGEDAPDLLKVITKPFTKLIVQLDK KVISQIAMNDEKAKNKSLVKIWCKTFTNKTQ INVTVPSTANCTSPSLCWTDGIQNWTMKNVT YKENIAKCQHIFVNFHLPDLAVGTILLILSLLV LCGCLIMIVKILGSVLKGQVATVIKKTINTDFP FPFAWLTGYLAILVGAGMTFIVQSSSVFTSAL TPLIGIGVITIERAYPLTLGSNIGTTTTAILAAL ASPGNALRSSLQIALCHFFFNISGILLWYPIPFT RLPIRMAKGLGNISAKYRWFAVFYLIFFFLIP LTVFGLSLAGWRVLVGVGVPVVFIIILVLCLR LLQSRCPRVLPKKLQNWNFLPLWMRSLKPW DAVVSKFTGCFQMRCCCCCRVCCRACCLLC GCPKCCRCSKCCEDLEEAQEGQDVPVKAPET FDNITISREAQGEVPASDSKTECTAL
1332	2682	A	10354	30	1377	SQQGSQPHRQGPPSLLTAPHSLDLPALPPGPR GSQGKLRRVLVPMSVKPSWGPGPSEGVTAVP TSDLGEIHNWTELLDLFNHTLSECHVELSQST KRVVLFALYLAMFVVGLVENLLVICVNWRG SGRAGLMNLYILNMAIADLGIVLSLPVWMLE VTLDYTWLWGSFSCRFTHYFYFVNMYSSIFF LVCLSVDRYVTLTSASPSWQRYQHRVRRAM CAGIWVLSAIIPLPEVVHIQLVEGPEPMCLFM APFETYSTWALAVALSTTILGFLLPFPLITVFN VLTACRLRQPGQPKSRRHCLLLCAYVAVFV MCWLPYHVTLLLTLHGTHISLHCHLVHLLY FFYDVIDCFSMLHCVINPILYNFLSPHFRGRLL NAVVHYLPKDQTKAGTCASSSSCSTQHSIIIT KGDSQPAAAAPHPEPSLSFQAHHLLPNTSPISP TQPLTPS
1333	2683	A	10358	2	884	AAGAGADGREPASERASRAEPPAVAMGQND LMGTAEDFADQFLRVTKQYLPHVARLCLIST FLEDGIRMWFQWSEQRDYIDTTWNCGYLLA SSFVFLNLLGQLTGCVLVLSRNFVQYACFGLF GIIALQTIAYSILWDLKFLMRNLALGGGLLLL LAESRSEGKSMFAGVPTMRESSPKQYMQLGG RVLLVLMFMTLLHFDASFFSIVQNIVGTALMI LVAIGFKTKLAALTLVVWLFAINVYFNAFWT IPVYKPMHDFLKYDFFQTMSVIGGLLLVVAL GPGGVSMDEKKKEW
1334	2684	A	10367	59	1562	QAWSLQVALSPFFFPASPSNSFAAAVPQLLFP ELPLPHVPGQESAKRRSARRFLIMSELTKELM ELVWGTKSSPGLSDTIFCRWTQGFVFSESEGS ALEQFEGGPCAVIAPVQAFLLKKLLFSSEKSS WRDCSQEEQKELLCHTLCDILESACCDHSGS YCLVSWLRGKTTEETASISGSPAESSCQVEHS SALAVEELGFERFHALIQKRSFRSLPELKDAV LDQYSMWGNKFGVLLFLYSVLLTKGIENIKN EIEDASEPLIDPVYGHGSQSLINLLLTGHAVSN VWDGDRECSGMKLLGIHEQAAVGFLTLMEA LRYCKVGSYLKISKIPYLDCLASETHLTVFFA KDMALVAPEAPSEQARRVFQTYDPEDNGFIP DSLLEDVMKALDLVSDPEYINLMKNKLDPEG

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion LGIILLGPFLQEFFPDQGSSGPESFTVYHYNGL KQSNYNEKVMYVEGTAVVMGFEDPMLQTD DTPIKRCLQTKWPYIELLWTTDRSPSLN
1335	2685	A	10375	82	2929	TRTKRRLGREKAMASPPRGWGCGELLPFML LGTLCEPGSGQIRYSMPEELDKGSFVGNIAKD LGLEPQELAERGVRIVSRGRTQLFALNPRSGS LVTAGRIDREELCAQSPLCVVNFNILVENKM KIYGVEVEIIDINDNFPRFRDEELKVKVNENA AAGTRLVLPFARDADVGVNSLRSYQLSSNLH FSLDVVSGTDGQKYPELVLEQPLDREKETVH DLLLTALDGGDPVLSGTTHIRVTVLDANDNA PLFTPSEYSVSVPENIPVGTRLLMLTATDPDE GINGKLTYSFRNEEEKISETFQLDSNLGEISTL QSLDYEESRFYLMEVVAQDGGALVASAKVV VTVQDVNDNAPEVILTSLTSSISEDCLPGTVIA LFSVHDGDSGENGEIACSIPRNLPFKLEKSVD NYYHLLTTRDLDREETSDYNITLTVMDHGTP PLSTESHIPLKVADVNDNPPNFPQASYSTSVT ENNPRGVSIFSVTAHDPDSGDNARVTYSLAE DTFQGAPLSSYVSINSDTGVLYALRSFDYEQL RDLQLWVTASDSGNPPLSSNVSLSLFVLDQN DNTPEILYPALPTDGSTGVELAPRSAEPGYLV TKVVAVDKDSGQNAWLSYRLLKASEPGLFA VGLHTGEVRTARALLDRDALKQSLVVAVED HGQPPLSATFTVTVAVADRIPDILADLGSIKTP IDPEDLDLTLYLVVAVAAVSCVFLAFVIVLLV LRLRRWHKSRLLQAEGSRLAGVPASHFVGV DGVRAFLQTYSHEVSLTADSRKSHLIFPQPNY ADTLLSEESCEKSEPLLMSDKVDANKEERRV QQAPPNTDWRFSQAQRPGTSGSQNGDDTGT WPNNQFDTEMLQAMILASASEAADGSSTLGG GAGTMGLSARYGPQFTLQHVLQGELGSDYR QNVYIPGSNATLTNAAGKRDGKAPAGGNGN KKKSGKKEKK
1336	2686	A	10379	I	557	RPRRRQPSFSCRVLVLEDPPCFRFTNSMNQEK LAKLQAQVRIGGKGTARRKKKVVHRTATAD DKKLQSSLKKLAVNNIAGIEEVNMIKDDGTVI HFNNPKVQASLSANTFAITGHAEAKPITEMLP GILSQLGADSLTSLRKLAEQFPRQVLDSKAPK PEDIDEEDDDVPDLVENFDEASKNEAN
1337	2687	A	10380		1263	PGSTISWSPAAARGLSVCRCCRLHPASAMDL FGDLPEPERSPRPAAGKEAQKGPLLFDDLPPA SSTDSGSGGPLLFDDLPPASSGDSGSLATSISQ MVKTEGKGAKRKTSEEEKNGSEELVEKKVC KASSVIFGLKGYVAERKGEREEMQDAHVILN DITEECRPPSSLITRVSYFAVFDGHGGIRASKF AAQNLHQNLIRKFPKGDVISVEKTVKRCLLD TFKHTDEEFLKQASSQKPAWKDGSTATCVLA VDNILYIANLGDSRAILCRYNEESQKHAALSL SKEHNPTQYEERMRIQKAGGNVRDGRVLGV LEVSRSIGDGQYKRCGVTSVPDIRRCQLTPND RFILLACDGLFKVFTPEEAVNFILSCLEDEKIQ TREGKSAADARYEAACNRLANKAVQRGSAD NVTVMVVRIGH
1338	2688	A	10385	3	589	GPSQSMAAGELEGGKPLSGLLNALAQDTFHG YPGITEELLRSQLYPEVPPEEFRPFLAKMRGIL KSIASADMDFNQLEAFLTAQTKKQGGITSDQ AAVISKFWKSHKTKIRESLMNQSRWNSGLRG LSWRVDGKSQSRHSAQIHTPVAIIELELGKYG QESEFLCLEFDEVKVNQILKTLSEVEESISTLIS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QPN
1339	2689	A	10386	50	390	LGAMAKHHPDLIFCRKQAGVAIGRLCEKCDG KCVICDSYVRPCTLVRICDECNYGSYQGRCVI CGGPGVSDAYYCKECTIQEKDRDGCPKIVNL GSSKTDLFYERKKYGFKKR
1340	2690		10388	-	3472	SQLRKGASATHSSPSRTDCIAQMMDIYVCLK RPSWMVDNKRMRTASNFQWLLSTFILLYLM NQVNSQKKGAPHDLKCVTNNLQVWNCSWK APSGTGRGTDYEVCIENRSRSCYQLEKTSIKIP ALSHGDYEITINSLHDFGSSTSKFTLNEQNVSL IPDTPEILNLSADFSTSTLYLKWNDRGSVFPHR SNVIWEIKVLRKESMELVKLVTHNTTLNGKD TLHHWSWASDMPLECAIHFVEIRCYIDNLHFS GLEEWSDWSPVKNISWIPDSQTKVFPQDKVIL VGSDITFCCVSQEKVLSALIGHTNCPLIHLDGE NVAIKIRNISVSASSGTNVVFTTEDNIFGTVIF AGYPPDTPQQLNCETHDLKEIICSWNPGRVTA LVGPRATSYTLVESFSGKYVRLKRAEAPTNES YQLLFQMLPNQEIYNFTLNAHNPLGRSQSTIL VNITEKVYPHTPTSFKVKDINSTAVKLSWHLP GNFAKINFLCEIEIKKSNSVQEQRNVTIKGVE NSSYLVALDKLNPYTLYTFRIRCSTETFWKW SKWSNKKQHLTTEASPSKGPDTWREWSSDG KNLIIYWKPLPINEANGKILSYNVSCSSDEETQ SLSEIPDPQHKAEIRLDKNDYIISVVAKNSVGS SPPSKIASMEIPNDDLKIEQVVGMGKGILLTW HYDPNMTCDYVIKWCNSSRSEPCLMDWRKV PSNSTETVIESDEFRPGIRYNFFLYGCRNQGY QLLRSMIGYIEELAPIVAPNFTVEDTSADSILV KWEDIPVEELRGFLRGYLFYFGKGERDTSKM RVLESGRSDIKVKNITDISQKTLRIADLQGKTS YHLVLRAYTDGGVGPEKSMYVVTKENSVGI IIAILIPVAVAVIVGVVTSILCYRKREWIKETFY PDIPNPENCKALQFQKSVCEGSSALKTLEMNP CTPNNVEVLETRSAFPKIEDTEIVSPVAERPEN RSDAKPENHVVESYCPPIIEEEIPNPAADETGG TAQVIYIDVQSMYQPQAKPEEEQENDPVGGA GYKPQMHLPINSTVEDIAAEEDLDKTAGYRP QANVNTWNLVSPDSPRSIDSNSEIVSFGSPCSI NSRQFLIPPKDEDSPKSNGGGWSFTNFFQNKP ND
1341	2691	A	10392	1	5057	MLPPKHLSATKPKKSWAPNLYELDSDLTKEP DVIIGEGPTDSEFFHQRFRNLIYVEFVGPRKTL IKLRNLCLDWLQPETRTKEEIIELLVLEQYLTII PEKLKPWVRAKKPENCEKLVTLLENYKEMY QPEGESLHGVLVVSAGLRCPLGLSASTLLTW SGLDNSLSWAAVGMSCVLWDIELHHDFLGV ATKSVSTHAQGDAAQGLGGTIVRMWARDSN LATGVLLDDNNSDVTSDDDMTRNRRESSPPH SVHSFSGDRDWDRRGRSRDTEPRDRWSHTR NPRSRMPPRDLSLPVVAKTSFEMDREDDRDS RAYESRSQDAESYQNVVDLAEDRKPHNTIQD NMENYRKLLSLGVQLAEDDGHSHMTQGHSS RSKRSAYPSTSRGLKTMPEAKKSTHRRGICED ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ RFPRMSDDNWKDISLNKRESVIQQRVYEGNA FRGGFRFNSTLVSRKRVLERKRYHFDTDGK GSIHDQKGCPRKKPFECGSEMRKAMSVSSLS SLSSPSFTESQPIDFGAMPYVCDECGRSFSVIS EFVEHQIMHTRENLYEYGESFIHSVAVSEVQK

NO. of much entitle sequence where the coation process of the coation of the coat	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide seq- uence ue							De Asportio Asid Es-Clutomio Asid
uence unce where and the service of the service of the service of peptide residue of peptide sequence where the sequence where the sequence where the sequence where the sequence where the sequence where the sequence where the sequence sequence where the sequence sequence where the sequence sequence where the sequence sequence sequence where the sequenc			1104	4			
uence 9,146 9,146 ng to first amino acid residue of peptide residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first amino acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue of peptide sequence 9,04 ng to first acid residue 9,04 ng to first acid residue 9,04 ng to first acid residue 9,04 ng to first acid residue 9,04 ng to first acid residue 9,04 ng to first acid residue 9,04 ng to first acid residue 9,04 ng to first ac							
uence 914 ng to first amino acid residue of peptide pepti							
amino acid residue of peptide requence peptide sequence peptide sequence peptide sequence peptide sequence per solo podon, peptide sequence per solo podon, peptide sequence per solo podon, peptide sequence per solo podon, peptide sequence per solo podon, peptide sequence per solo podon, peptide sequence per solo per		uence	1				
residue of peptide sequence Y=Tyrosine, X=Unknown, **Siop codon, /possible mucleotide letion, \possible mucleotide letion, \possible mucleotide linertion SVGGKRFECKLOCHTFINSAALABERKIHA RÖYLVECKNOGEEAPMPSTTSELQKIYGK SVEYGKEVCKETHI-HSSALEHGKIHGODKD NEREHRERERERGETFRFSPALNFFCKMYG KEKMYECKVCGETFINSSALEHGKIHGODKD NEREHRERERERGETFRFSPALNFFCKMYG KEKMYECKVCGETFINSSALEHGKIHGODKD NEREHRERERGEETFRFSPALNFFCKMYG KEKMYECKVCGETFINSSALEHGKHITGON PFENKGKVCEGETFINGSSLERKGKTYNKEELC DFTDGRAPMQSSELSHOKHISKNIFEGR GYEKSVIHSIGPTESOSLERKGKTYNKEELC DFTDGRAPMQSSELSHOKHISKNIFEGR GYEKSVIHSIGPTESOSLERKGKTYNKEELC DFTDGRAPMGSSELSHOKHISKNIFEGR GYEKSVIHSIGPTESOSLERKGKTYNKEELC DFTDGRAPMGSSELSHOKHISKNIFEGR GYEKSVIHSIGPTESOSLERKGKTYNKEELC DFTDGRAPMGSSELSHOKHISKNIFEGR GYEKSVIHSIGSVAGSPKYKWAASSINGSTPAKYGKSVY GEGSGEFKKDGEFSVESSVIHSISSELSHOKHISKNIFEGR GYEKSVIHSIGPTSSSVIHSISSELSHOKHISKNIFEGR GYEKSVIHSIGPTSSSVIHSTSVIHSISPTSVIHSIPTSDEDEKA APPOELSESINGSVINGSVFAKYGKSVY GEGSGEFKKDGEFSVESSVIHTSISSEVGARINCK DFRQFFATSEDLINDKROKUPAR GEGSGEFKKDGEFSVESSVIHTSISSEVGARINCK DFRQFFATSEDLINDKROKUPAR GEGSGEFKKDGEFSVESSVIHTSISSEVGARINCK DFRQFFATSEDLINDKROKUPAR GEGSGEFKKDGEFSVESSVIHTSISSEVGARINCK DFRQFFATSEDLINDKROKUPAR GEGSGEFKKDGEFSVESSVIHTSISSEVGARINCK DFRQFFATSEDLINDKROKUPAR GEGSGEFKKDGEFSVESSVIHTSISSEVGARINCK DFRQFFATSEDLINDKROKUPAR GEGSGEFKKDGEFSVESSVIHTSISSEVGARINCK DFRQFFATSEDLINDKROKUPAR GEGSGEFKKDGEFSVESSVIHTSISSEVGARINCK DFRQFFATSEDLINDKROKUPAR GEGSGEFKKDGEFSVESSVIHTSISSEVGARINCK DFRQFFATSEDLINDKROKUPAR GEGSGEFKKDGEFSVESSVIHTSISSEVGARINCK DFRQFFATSEDLINDKROKUPAR GEGSGEFKKDGEFTSSVESSVIHTSISSEVGARINCK GEGSGEFKKDGEFTSSVESSVIHTSISSEVGARINCK GEGSGEFKKDGEFTSSVESSVIHTSISSEVGARINCK GEGSGEFKKDGEFTSSVESSVIHTSISSEVGARINCK GEGSGEFKKDGEFTSSVESSVIHTSISSEVGARINCK GEGSGEFKKDGEFTSSVESSVIHTSISSEVGARINCK GEGSGEFKKDGEFTSSVESSVIHTSISSEVGARINCK GEGSGEFKKDGEFTSSVESSVESSVESSVESSVESSVESSVESSVESSVESSV	uence		ļ	914		1	
peptide sequence Sequence	1				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
peptide			1	ł l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
	[ĺ	1	peptide	-	/=possible nucleotide deletion. \=possible
SOVGGKFERKDGGETFINSSALLAEIRKIHA RGYLVECKNOGCEBAPMPSTISELGKIYGK DKFYECRVCKETFLHSSALLEIDKIHGDDKD DKFYECRVCKETFLHSSALLEIDKIHGDDKD NEREIRREERERGEGGETFFSPSALNEFOKMYO KEKMYECKVCGETFIGSSLKEHOKHTROM PEPKKKOVCETFIGGSLKEROKYNKEELC DPTDGRDAPMOSSELSEHOKHSRKNLFOR GYEKSYHISGPTESOKSHTITRILESDEDEKA FTISSNPYENOKIPTKENVYRAKSYERSVHISL AVVRAOKSISVAGPSKYMAESTIGSEDAN HORVRAGGINTSEGREYSRSVHISL VASKPPRS HIGHELVESNEKGESSINTSDLINDKROKIPAR ENPCEGGSKINNYEDSVIOSVPRAKPOKSIPA HORVRAGGINTSEGREYSRSVHISLASKPRS HORVELVESNEKGESSINTSDLINDKROKIPAR ENPCEGGSKINNYEDSVIOSVPRAKPOKSVPA HORVELVESNEKGESSINTSDLINDKROKIPAR ENPCEGGSKINNYEDSVIOSVPRAKPOKSVPA HORVELVESNEKGESSINTSDLINDKROKIPAR HORVRAGGINTSEGREYSRSVHISLASKPRS HORVELVESNEKGESSINTSDLINDKROKIPAR HORVRAGGINTSEGREYSRSVHISLASKPRS HORVELVESNEKGESSINTSDLINDKROKIPAR HORVRAGGINTSEGREYSRSVHISLASKPRS HORVELVESNIKEVENSTALLESSINTSDLINDKROKIPAR HORVELVESNIKEVENSTALLESSINTSDLINDKROKIPAR HORVRAGGINTSEGREYSRSVHISLASKPRS HORVELVESNIKEVENSTALLESSINTSDLINDKROKIPAR HORVELVESNIKEVENSTALLESSINTSDLINDKROKIPAR HORVELVESNIKEVENSTALLESSINTSDLINDKROKIPAR KOCDDGIALTSEGLEDGUTTERVOK SKICLVOSSETTHISVIHTHSISEVARALESSINTSDLINDKROKIPAR KOCDDGIALTSEGLEDGUTTERVOK KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDGUTTER KOCDDGIALTSEGLEDG	1					ļ	nucleotide insertion
RÖYLVECKNÖEGERAPMSPITSELQKIYGK DEFYECVCKETHEISALEHGKIHGODKD NERHERERERERGETFRSSALEHGKITGKOMYG REKMYECKVCGETHISSALEHGKIHTGON PENKGKVCEGETFRGGSLKRRQKIYNKERLC DPTOGRAPMQSSELSEHGKIHSTKNIJEGR GVEKSVIHIGGFTESQKSHTITRILESDEDEKA FITISSPYENQKIPTKENVYEAKSYERSVIHSI ASVEAQKSHSVAGFSKRKVMAESTIGSEDARM HORVRAGGNTSGERGKSYRSVIHSILVASKPRS RIGORLVESNEKGESSIYISDLINKRQKIPAR BPNCEGGSKRINYEDSVIJQSYFRAKPQKSVP GGSGEFKADGESPYSSSNYREVQKARAKKK YIBHSNETSVIHSI.PFGGOTFRRGMI, YECQ EGGCAHSSDLTEHQKHENEKSPSSRNYTE WSVIRSLAPTIDQTSYAGEQYAKEQARINKCK DPROFATSEDLINTINGKIYDQEKSIGGESOGE NITOGETHISETHIGQETTIEDPVIQGSDMEDPQ KDDPDDSTYECEDCGIGFVDLIDLITDHQKVH SRKCLVDSREYTHSVHTHSISEQQRDYTGEQ LYECPKCGSENHSSTLEHGPKHEQDQLYSM KGCDDGFIALLPMKPRRNRAAERNPALAGSA BRCLLCGGFHISSALHWINHREDPLEQS QMAEEAIIPGLALTEFGRSQTEERLFECAVCG ESFVNPAELADHTVYTHKINEPYEVGSSYTHTS FLTEPLKGAIFFYECKDCGSSPHISTVLTKIKEE LHLEEEEEDBAAAAAAAAQVEVANWIVPQ VVLRIQGLNVEAAEPEVAAEPEVEAEPEVEAEPEVEAEPEVEAEPEVEAEPEVEAEPEVEAEPEVEAEPEVEAEPEVEAEPEVEAEPEVEAEPEVEAEPEVEAEPEVE	<u> </u>				Sequence		
DKFYECRVCKEITH.HSSALEHQKIHGDDKD NREHRERERERERFERSPANLEFCKMYQ KEKMYECKVCGETTEH.SSSLKEHQKIHTRGN PERMKGVCECTTERGSQLKERQKTYNKEELLC DFTDGRDAPMQSSELSEHQKIHSRKNILEGR GYEKSVHISGPTESQGSLKERQKTYNKEELLC DFTDGRDAPMQSSELSEHQKIHSRKNILEGR GYEKSVHISGPTESQGSLKENQKTYNKEELLC DFTDGRDAPMQSSELSEHQKIHSRKNILEGR GYEKSVHISGPTESQGSLKHTIRPLESDEDEKA PTISSNPYENQKIPTKENDVPGAKSYERSVHISL AVVEAQKSHSVAGPSKIPKVMAESTIQSFDAN HQRVRAGGNISEGRETSRSVHISLVASSPRS HMGNELVESNEKGESSIYISDLINDKRCKIPAR ENPEGGGSKRINNYBDSVIQSVFRAKPGKSVP GGGSGFKKDGEFSVPSSNVREVQKARAKKK YEHERINSTSVHISLPFGGTFRPRGNTYECQ ECGCTAHSSDLTEHQKHDREKPSGSRNYE WSVIRSLAPIPOPTSYAQEQYAKEQARNICK DFRQFFATSEDLNTNQKIYQQEKSIGGESQGE NTDGGEHSBETHGGTEDPVIQGSDMEDPQ KDDPDDKIYECEDCGLGFVDLITDLITDHQKVH SRKCLVDSREYTHSVHTHSISEVQRAPINGC ULYCPKCGSSPHISSFLFEHQRHEDDOLLOS SRKCLVDSREYTHSVHTHSISEVQRAPINGCA ULYCPKCGSSPHISSFLFEHQRHEDDOLLOS SRKCLVGGGFIHSSALNEHMRLHREDDLLEQS GAREARIPGLALTEFROSTGEELIPECAVCG ESFYNPAELADHVIVHKNPPVYGSSYTHTS FLFEPKGAPFYEKCHOCGSSPHISTVITKHKE LHEEGEEDBAAAAAAAAAAQVEANVHVPQ VURKIGGINVEAAREPECKAGEPE GDADEPDGAGIEDPEGAGEPEGAGQNGGAB QONGDADEPDGAGIEDPEGAGEPEGAGQNGGAB GONGDADEPDGAGIEDPEGAGEPEGAGQNGGAB GCSGYTERASTSTGGANQADEKYTRCDVITKHKE LHEEGEEDBAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	J		j	j .		,	
NEREHRERERGETREPSALNEFOKHTRON FENKGKVCEVTGETI-LISSIS,KERIKTRON FENKGKVCEVTGETI-LISSIS,KERIKTRON FENKGKVCETTFIPGOSLKRROKTTYKKELL DFITDGRAPMOSSESJEHQKIBERNIL-FEGR GYEKSVHISGPTESQKSHTTRPLESDEDEAK ATISNSPYENOKUPTKENV-YEAKSYERSVHISL ASVEAQKSHSVAGPSKPK VMAESTIQSEDAIN HQR VERAGGINTEGREYSSEVSHI ASKEPRS HNONEL-VESNEKGESSIYISDI.NDKROKUPAR ENPCEGGSKIRNYNEDBVIQSVPRAPKOKYV GEGGSCEPKKDGEFSVPSSNVREYQKARAKKK YIERISNETSVIRLSIPFOGGTPRPROMLYECQ ECGECFAHSDLTEHQKIHDRKPSGSNYYE WSTRELATTPOPTSTAQEQYAKEQARNKCX DFRQFFATSEDLNTNQKITQDEKSPGSNYYE WSTRELATTPOPTSTAQEQYAKEQARNKCX DFRQFFATSEDLNTNQKITQDEKSGSGSGE NTDGETHSEETHQGETIEDPVIQGSDMEDDO KDPPDDKTPEEDCOLGFVDLTDTHQKVH SRKCLVDSREYTHSVHITHSISEYQRDYTEGE LIVECPKGESSHISSFLEBHQRIHEQDQL'VSM KGCDDGFALLPMKPRRNRAAERNPALAGS RCLLCGQGFHSSALINEHMRLHEDDLLGSQ QMAEAJIBGLALTEPQRSQTERLFECAVCO ESFVNPAELADHVYHKNEPEYEYGSSYTHTS FLITEPLKGAJFFYECKDCGKSFHSTVLTKHER LILLEEEEDEAAAAAAAAAGPEVVEAAEPEV EAAEPNGEASEPDGEAAPFIGEAQPNIGEAE QPNGDADEPDGAGEDEFERAEEPFGKAEPEY EAAEPNGEASEPDGEAAPFIGEAQPNIGEAE QPNGDADEPDGAGEDPEFERAEEPFGKAEPEY EAAEPNGEASEPDFEGAAPFIGEAGPFYDC HECTETTTSSTAFSEHLKTHASMIIFEPANATHY VVLRIQGLNVEAAFEVVEAAEPEVYBACPEYDO HECTETTTSSTAFSEHLKTHASMIIFEPANATHY PGSSEVPKLINTVQDQEEBRCEGGAJSLIQ ECGGYBRASTSTGAAQABEKYFKCDVCGG ECGGYBRASTSTGAANQABEKYFKCDVCGG LYNDHLSLARHQNTHIG GERRSSSDNNRIVERGAGLISSAAVQTRIGNSA ASRRSPAARPPYPAPALFRGPGTGGSTSL VCNPAIFELLARQLAKHAHMSWIPESAD AVLVVAVAVVVVSVAVVAVAMANNTHY PGSSEVPKLINTVQDQEEBRCREGGALDPKH VCNPAIFELLARQLAKHAHMSWIPESAD UGEENTWAKELSDLASSVALCHNDLLCKNII YNEKQGDVQFIDYEVSGYNVLCHNDLLCKNII YNEKQGDVQFIDYEVSGYNVLATDIONISPH PAGYSDVDYSLIPPGALDPKHAYERW ERGGGTTUFFGADEDINKRPLSDIPSSQI LQEENTWAKELSDLASSVALCHNDLLCKNII YNEKQGDVQFIDYEVSGYNVLATDIONISPH PAGYSDVDYSLIPPGALDPKHAYERW ERGGGTETTERGOPHLGCQCGG UMPTRESSTANAVQDWSTORLTYSHYMPK ERGGGTETTERGOPHLGCQCGG LMYDDINGESTRETERGOPHLGCQCGG FFFSTSAKDNINVXQCTFFERLVDILCKMESLEF FFSTSAKDNINVXQCTFFERLVDILCKMESLEF FFSTSAKDNINVXQCTFFERLVDILCKMESLEF FFSTSAKDNINVXQCTFFERLVDILCKMESLEF FFSTSAKDNINVXQCTFFERLVDILCKMESLEF FFSTSAKDNINVXQCTFFERLVDILCKMESLEF FFSTSAKDNINVXQCTFFERLVDILCKMESLEF	İ				i '		
KEKMYECKVCGETTLHSSILGENGKITIKERLO PFENKGKVCEETPIGOSLKRROKTYNKERLC DFTDGRDAPMQSSELSEHQKIHSRKNIPEGE GYEKSVHSOPTPIGOSLKRROKTYNKERLC DFTDGRDAPMQSSELSEHQKIHSRKNIPEGE GYEKSVHSOPTPISQSKITITIPHESDEDEKA FTISSNPYENOKIPTKENVYBAKSYERSVIHSI ASVEAQKSHSVAGPKSPK WALESTIQSPDAN HORVRAGONTSEGREYSRSVHISLVASKPRS HIMONELVESNEKGESSIYISDLNDKRQKIPAR EPPECEGSKRRNYDEDSVIQSVRAKFYGKSVP GEGSGEFKKDGEFSVPSSNNREYQKARKKK YIEHRSNETSVHISLIPFGEQTTEPRYQKAPKKKK YIEHRSNETSVHISLIPFGEQTTEPRYQKAPKKKK JYEHRSNETSVHISLIPFGEQTTEPRYQKSMHYMEQ ECGECFAHSSDLTEHQKHIDHEKPSGRNYPY WSWIRSLAPTDPGTYSAQEQYAKEYGKNYK GEDSGEFKLSTEHQKHIHGQDLYSW KOLDPDDKYTSCEDCGLGFVDLTDLTDHOKYK SKCLVDSREYTHAVHTHSISEYQRDYTGEQ LYECPKGESFHSSFLFEHQRIHRDQDLYSM KGCDDGFIALIPMKYRRNRAARENPALAGSA IRCLLCGQGFHISSALNEHMRI-HREDDLLGS QMAEAIRFGALTHEROPSOLTSH KGCDDGFIALIPMKYRRNRAARENPALAGSA IRCLLCGQGFHISSALNEHMRI-HREDDLLGSQ QMAEAIRFGALTHEROPSOLTSH KGCDDGFALLTBKYRRNRAARENPALAGSA IRCLLCGQGFHISSALNEHMRI-HREDDLLGSQ QMAEAIRFGALTHEROPSOLTSH KGCDDGFALLTBKYRRNRAARENPALAGSA IRCLLCGQGFHISSALNEHMRI-HREDDLLGSQ QMAEAIRFGALTHEROPSOLTSH KGCDDGFALLTBKATTARTORTHEROPSOLTSH LIFLERGERDAAAAAAAAAQEVEANVHIVPQ VVLRIGGLINGBAAGPEVEAAEPFYGGSYTTHTS FLTEPLKGAIPFYECKDCKGSFHISTVLTKIKKE LILLEEBEEDBAAAAAAAAQEVEANVHIVPQ VVLRIGGLINGBAAGPEVEAAEPFYGGSVCACHEPE GDADEPDGVGIEDPEGEAGEPFGGSATSH QPNGGADAPPTOGAIDPERAEPFYEAAEPFYCOCCG ESFVIPAELADHYTVHKNEPYETGOSYTTHT FLTEPLKGAIPFYECKDCKGSFHISTVLTKIKKE LILLEEBEEDBAAAAAAAAAQEVEANVHIVPQ VVLRIGGLINGBAAGPEVEAAEPFYTOCCHORY FLTERGAIPTACHAPPTYTOCCHORY FLTER				1		j -	
PFENKGKVCETTPIGOSLKRRQKTYNKELG DFTDGRAPMQSSELSEHQKIRRINLFEGR GYEKSVHISQPTESQKSHITIRPLESDDEJAK HORDRAPMQSSELSEHQKIRRINLFEGR GYEKSVHISQPTESQKSHITIRPLESDDEJAK ASVEAQKSHSVAGPEKPKVMAESTIQSEDAN HQRVRAGONTSEGREYSRSVHISLVASKPRS HIMONELVESNEKGESSIYISDLNDKRQKIPAR ENPECGGSKRRNYEDSVIQSVFRAKPQKSVY GEGSGEKKDGEFSVPSSNVREYQKARAKK YIEHRSNETSVHISLIPGEQITERPRQMLYECQ EGGECFAHSSDLTEHQKHIDERKSPGSRYYE WSVIRSLAPTDPQTSVAQEQYAKEQARNKC DPRQFFATSEDLNINQKIYDQEKSIGEESQGE NTDGETHISETHQGETIEDPVIQGSDMEDPQ KDPPDDKIYSECDEGGGFVLDTIDHQKVH SRKCLVDSRSYTHSVHITHSISEYQRDYTGGC LYCEPKGGSEFHASSLEHGMIRHEDQLYSM KGCDDGFALLPMKPRRNRAAERNPALAGSA RCLLCQGGHISSALNEHMIRHEDQLYSM KGCDDGFALLPMKPRRNRAAERNPALAGSA RCLLCQGGHISSALNEHMIRHEDQLYSM KGCDDGFALLPMKPRRNRAAERNPALAGSA RCLLCQGGHISSALNEHMIRHEDQLYSM KGCDDGFALLPMKPRRNRAAERNPALAGSA RCLLCQGGHISSALNEHMIRHEDQLYSM VALRIQGLNVEAAFPEVEAAEPYEVEAPPYEV EAAEPNGEAGPDGEAAFPEVEAAEPPE GDADEPDOVGEDPEGGEDGEQVERNYTVD VVLRIQGLNVEAAFPEVEAAEPPEVAAEPPE GDADEPDOVGEDPEGGEDGEQVERNYTVD HECTETTTSSTATSBLIKTHASMIFPENAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LYNDHILLARHQNITHG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLERGGISSAAVQTRIGNSA ASRRSPAAPPPVPAPALPRGRPGTGGSTSIS.A ARAVLVVAVAVAVVVVVSAVAWAMANNYTHY PPGGSPEVPLINYTVQDQEEHRCREGALSLU, QAHGGACPQLYCTPNOLCYETIQGEALDPKH VCNPAFFILLARQLAKHLAHANNGWPKSNL WKKMGYTSLIFTGADDINKRFLSIGPSSQU LQEEMTWMKELSNLGSVVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLCKNII VPEKGGDVGFIDYSTGNVLCHNDLCKNII VPEKGGDVGFIDYSTGNVLCHNDLCKNII VPEKGGDVGFIDYSTGNVLCHNDLCKNII VPEKGGDVGFIDYSTGNVLCHNDLCKMSSLJET DPATIAAKONTLKETPPPOPONCAC DPPWWSKVDDPVTNLIHLSSSKGRGH	1			į ,		1	NEREHERERERERGETFRPSPALNEFQKMYG
PFENKGKVCETTPIGOSLKRRQKTYNKELG DFTDGRAPMQSSELSEHQKIRRINLFEGR GYEKSVHISQPTESQKSHITIRPLESDDEJAK HORDRAPMQSSELSEHQKIRRINLFEGR GYEKSVHISQPTESQKSHITIRPLESDDEJAK ASVEAQKSHSVAGPEKPKVMAESTIQSEDAN HQRVRAGONTSEGREYSRSVHISLVASKPRS HIMONELVESNEKGESSIYISDLNDKRQKIPAR ENPECGGSKRRNYEDSVIQSVFRAKPQKSVY GEGSGEKKDGEFSVPSSNVREYQKARAKK YIEHRSNETSVHISLIPGEQITERPRQMLYECQ EGGECFAHSSDLTEHQKHIDERKSPGSRYYE WSVIRSLAPTDPQTSVAQEQYAKEQARNKC DPRQFFATSEDLNINQKIYDQEKSIGEESQGE NTDGETHISETHQGETIEDPVIQGSDMEDPQ KDPPDDKIYSECDEGGGFVLDTIDHQKVH SRKCLVDSRSYTHSVHITHSISEYQRDYTGGC LYCEPKGGSEFHASSLEHGMIRHEDQLYSM KGCDDGFALLPMKPRRNRAAERNPALAGSA RCLLCQGGHISSALNEHMIRHEDQLYSM KGCDDGFALLPMKPRRNRAAERNPALAGSA RCLLCQGGHISSALNEHMIRHEDQLYSM KGCDDGFALLPMKPRRNRAAERNPALAGSA RCLLCQGGHISSALNEHMIRHEDQLYSM KGCDDGFALLPMKPRRNRAAERNPALAGSA RCLLCQGGHISSALNEHMIRHEDQLYSM VALRIQGLNVEAAFPEVEAAEPYEVEAPPYEV EAAEPNGEAGPDGEAAFPEVEAAEPPE GDADEPDOVGEDPEGGEDGEQVERNYTVD VVLRIQGLNVEAAFPEVEAAEPPEVAAEPPE GDADEPDOVGEDPEGGEDGEQVERNYTVD HECTETTTSSTATSBLIKTHASMIFPENAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LYNDHILLARHQNITHG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLERGGISSAAVQTRIGNSA ASRRSPAAPPPVPAPALPRGRPGTGGSTSIS.A ARAVLVVAVAVAVVVVVSAVAWAMANNYTHY PPGGSPEVPLINYTVQDQEEHRCREGALSLU, QAHGGACPQLYCTPNOLCYETIQGEALDPKH VCNPAFFILLARQLAKHLAHANNGWPKSNL WKKMGYTSLIFTGADDINKRFLSIGPSSQU LQEEMTWMKELSNLGSVVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLCKNII VPEKGGDVGFIDYSTSGNVLCHNDLCKNII VPEKGGDVGFIDYSTGNVLCHNDLCKNII VPEKGGDVGFIDYSTGNVLCHNDLCKNII VPEKGGDVGFIDYSTGNVLCHNDLCKNII VPEKGGDVGFIDYSTGNVLCHNDLCKMSSLJET DPATIAAKONTLKETPPPOPONCAC DPPWWSKVDDPVTNLIHLSSSKGRGH					•	į .	KEKMYECKVCGETFLHSSSLKEHOKIHTRGN
DFTDGRDAPMQSSELSEHQUISRRNLIFEG GYEKSYHRISOPTESQKSHITTIPLESDEDEKA FTISSNPYENOKIPTKENVYPAKSYERSYHRIS ASVEAQKSHSVAGPSKYPKAMESTIGSFDAIN HQRVRAGGNTSEGREYSRSVIRISLVASKEPGANIN HQRVRAGGNTSEGREYSRSVIRISLVASKEPGANIN HQRVRAGGNTSEGREYSRSVIRISLVASKEPGANIN ENPCEGGSKYRNPEDBVIQSVFRAKPQKSVY GEGSGEPKKDGEFSVYPSSNVREYQKARAKKK YTHERSNETSYHISLPFGGGTTEPKROMLYEQ ECGECFAHSSDLTEHQKILHDRKPSGSRNYL WSVIRSLAPTIPOTIS VAQEQVARKKK DFRQFFATSEDLNTINQKIVTQGKSMEEDPQ KDDPPDKIVECEDCGLGFVDLTDLTDHQKYN SKCLVDSREYTHSVHITTISSEYQRDYTGEQ LVECPKGGSSFHISSILFEHQRIHEQDQLVSRY RKCLVDSREYTHSVHITTISSEYQRDYTGEQ LVECPKGGSFHISSILFEHQRIHEQDQLVSRY KGCDDGFIALIPMKPRRNRAARENPALAGSA IRCLLCQQFHISSALNEHMRHREDDLLGS QMAEALIPGLALTEPGRSQTEERHJECAVCG ESFVNPAELADHVTVHKNEPYEVGSSYTHTIS FLTEPLKGAIPFYECKDCKGSFHISTVLTKHKE LHLEEEEEDBAAAAAAAAQEVEANVHVYO VVLRIQGLNVEAAFEPEVAAEPEVEAPEPV EAAEPNGBAEGPGBAAPGBEAQGPNGEA QPNGGDAEPDGAGEDPERAEEPGKAEPEV GDADEPDGVGIEDPERGEDQEIQVEEPYYDC HECTETTISTSTAFSEHLKTHASMIIFEPANAFG ECSGYTERASTSTGGANQADEETYFCDVCGQ LYRDHILLARRONTHTG GRPRSSDNRNFLERRAGISSAAVQTRIGNSA ASRRSAARPVPAPAPALPRGRPGTEGSTSLIA APAVLVVAVAVVVVSAVAVVAMANNYHY PPGSFEVRKLNTVQDQEBRRCEGGALSLIQ HLRPHRUPDQEVTLIQFTDGTTIKHLIGCYVGN TMEDVVLVNIYGRKTELLVORDEEVKSFRVL QAHGGAPQLVCTFONIOLCYETGEGALDPKH VCNPAIFRLIARQLAKHAHHANGWIPKSNI VPKKGGDVGFTPYSGNOWAAMANNYHH PPGSFEVRKLNTVQDQEBRRCEGGALSLIQ UKUMLIQAKVSTEEPDFLGYAUFRHQYFKM KFEYTALKVPB GLWALIQAKVSTEEPDFLGYAIVRFNQYFKM KFEYTALKVPB GLWALIGAKVSTEEPDFLGYAIVRFNQYFKM KFEYTALKVPB GLWALIGAKVSTEEPDFLGYAIVRFNQYFKM KFEYTALKVPB GLWALIGAKVSTEEPDFLGYAIVRFNQYFKM KFEYTALKVPB GLWALIGAKVSTEEPDFLGYAIVRFNQYFKM KFEYTALKVPB GLWALIGAKVSTEEPDFLGYAIVRFNQYFKM KFEYTALKVPB GLWALIGAKVSTEEPDFLGYAIVRFNQYFKM KFEYTALKVPB GLWALIGAKVSTEEPDFLGYAIVRFNQYFKM KFEYTALKVPB GLWALIGAKVSTEEPDFLGYAIVRFNQYFKM KFEYTALKVPB FEAGYSDVDYSLYPDRELGSGWLTATILEAVSKGENGFKTOTYS FFETSAKDNINVKQTFFERLVDILICKSKGEKGHSPAKDTS 1344 2694 A 10395 2 4136 DEPPWRSVEDPTVILLILESSKGHISPAKDTS					ĺ	ľ	
GYEKSVIHSOPTESOKSHTTRPLESDEDEKA PTISSNPYENGKIPTKENVYAKSYERSVIHSL ASVEAQKSHSVAGPEKRVYAMESTIGSFADIN HQRVRAGGINTSEGREYSRSVIHSLVASKPPRS HNGNELVENIKGESSIYISDLADKIQKUPAR ENPCEGSKNRNYEDSVIQSVFRARPQKSVP GEGSGEFKKOGEFSVPSSNVERYQKARAKKK YIEHRSNETSVIHSLPTGEQTTRRIGMLYEQ ECGCFAHSSDLTEHQKHIDEKRPSGSRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK DFRQFFATSEDLATINQKYTQDEKSHGEESQGE NTDGEETHSEETHQGETIEDPVQGSDMEDPQ KDDPDKIYTECEDCGLGFFVLDTHDHQKVH SRRCLVDSREVTHSVHITHSISEYQRDYTGEQ LYECKCGESFHSSFLFEHQKHEQDQLYSM KGCDDGFALLPMRPRRNRARERNPALAGSA IRCILCGQGFHISSALNEHMERPDLLEGS QMAEEAIPGLALTEFQSQTEERLFCAVCG ESFVPABLADHVTVHCNEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFHISTVLTKKKE LHLEEEEDBAAAAAAAAQGVEANVIPQ VVIRQGINVEAAEPEVEAAEPEVGAAEPEV EAAEPNGBAEGPDGEAAEPIGGACQPNGCEA QPNGDADEPDGGGEDPEEGEQEQVEPTYDC HECTETTSTSTAFSEHLKITHASMIFEPANAPG ECSGYIERASTSTGGANQADEKYPKCDVCGQ LYDDHLSLARMQNTHTG GCGGGGGGDEFERAFGGGAEPE GDADEPDGVGIEDPEEGEQEQUEPYYDC HECTETTSTSTAFSEHLKITHASMIFEPANAPG ECSGYIERASTSTGGANQADEKYPKCDVCGQ LYDDHLSLARMQNTHTG GCGGGGGGDEFERAFGGGAEPE GDADEPDGVGIEDPEEGEQEQVEEPYYDC HECTETTSTSTAFSEHLKITHASMIFEPANAPG ECSGYIERASTSTGGANQADEKYPKCDVCGQ LYDDHLSLARMQNTHTG PPGSFEYKLINVTVQDQEEHRCGALSILQ HLRPHWDPQEVTLQLFIDGITNKLIGCYVGN TIMEDVALVRIVYGNOTELLVDRDEEVKSRVL QARIGCAPQLYCFTRONGLCYEFQGEALDPEH VCNPABTALIARQLAKHAHAHANGWIFKSNL WIKMGKYTSLIPTGADEDINKERISDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYSYGGYNVLAYDLAYK EKGGDVQFIDYSYGGYNVLAYDLAYK EKGGDVQFIDYSYGGYNVLAYDLAYK EKGGDVQFIDYSYGGYNVLAYDLAYK EKGGDVGTINYSGGYNVLAYDLAYK EKGGDVGTINYSGGYNVLAYDLAYK EKGGDVGTINYSGGYNVLAYLLAYK EKGGDVGTINYSGGYNVLAYLLAYK EKGGDVGTINYSGGYNVLAYLLAYK EKGGBVGTINYSGFORALSHIFFW GLWALJQAKYSTEFDFLGYALVFRQVYKM EKRIKLQIWDTAGQERYRITITAYYRAMGFI LMYDITNEESBNAVQDWSTJQDFKVKTVFKN EKRIKLQIWDTAGQERYRITITAYYRAMGFI LMYDITNEESBNAVQDWSTJQDFKVKTVFKN EKRIKLQIWDTAGQERYRITITAYYRAMGFI LMYDITNEESBNAVQDWSTJQDFKVKTVYKN EKRIKLQIWDTAGQERYRITITAYYRAMGFI LMYDITNEESBNAVQDWSTJQDFKVKTVYKN EKRIKLQIWDTAGQERYRITITAYYRAMGFI LMYDITNEESBNAVQDWSTJQDFLOKMSESLET DPAITAAKQNTKVGTFFEDPLOCICGFE FFETSAKDNINVKQTFERLVDDICKMSESLET DPAITAAKQNTKVGTFFEDPLOCICGF				į .	!	ľ	
FITSSNPYENQKIPTKENVYEAKSYERSYHISL ASVEAQKSHSVAGESREVAMESTIGSPDAIN HQRVRAGGNTSEGREYSRSVHISLVAKSPRS HNGNELVESNEKGESSITYSIDADKRQKIPAR ENPCEGGEKNRNYEDSVIGSVFRAKPQKSVP GEGSGFFKKDGEFSVPSSNVEBYQRAKKK YEHRISNETSVHISLPFEGGTIPTRIGMLYECQ ECGECFAHSSDLTEHQKHIDEKEPGSGRYYE WSVIRSLAPIPOGTSVAGEQARNKCK DFRQFFATSEDLNTNQKIYDQEKSHGEBGGE NTDGEETHISSETHQGTIETEPVQGSDMEDPQ KDPDDKIYECEDGGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSVHTHISLPFGGUTHQGSDMEDPQ KDDPDKIYECEDGGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSVHTHISLPFGQDLYSM KGCDGFIALIPMKPRRINAENPALAGSA IRCLLCQGFHISSALNEHMRLHREDDLLEQS QMAEEAIPGLALTEPGRSQTEERLFCAVCG ESFNYPAELADHVTVHKNEPYGSSYTHTS FITERLKGAIPFYECKDCGKSFHISTVLTKHKE LHLEEEEDBAAAAAAAAQOVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVAAEPEV EAAEPNGEAEGPDGEAAEPFOGAAOPPOGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGEAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGEAE QPNGDADEPDGWGIEDPEGGAGOPNGAAPG ARROSSDORDATAMANGWFSSUL WLKMGKYSLPTGPADATAAVGWFWGNLAWAA VLVVAVAVVVVVVAVAVAWAMANANININ YNEKQGDVQFIDYSTSTGRQHAGHAAH VCNPAFFELIADBATAAVGNWGIYGTYSWDNAQ QARYGGKDSSDONPOTYMGHASHCILIONSVGK TSFLFRYADDSFTSAFVSTVGIDFKWKTVFKN EKRILQIWDTAQGERYRTITTAYAYGMGFI LMTDITNESSRNAVQDWSTQIKTYSWDNAQ V	ļ			ĺ	,	ļ	GVEVEVIUCCDETECOVCHTITADI CODEDENA
ASVEAQKSHSVAGESREYNAMESTIGSPADN HQRVRAGGINTSEGREYSRSVUISLVASKPDRS HNGNELVESNEGESSIYISDLADKRQKIPAR ENPCEGSKNRNYEDSVIQSVFRARPQKSVP GEGSGEFKKOGEFSVPSSNVEYQKARAKKK YEHRSNETSVHSLPPGEQTFRYRGMLYEQQ ECGCFAHSSDLTENQKHIDEKPSGSRNYE WSVIRSLAPTDPQTSVAQEQVAKEQARNKCK DFRQFFATSEDLNTNOKIYUPQKSHIGEESQGE NTDGEETHSEETHIQETTEDPVIQGSDMEDPQ KDDPDKYYECEDCGLGFFVDLTDHOKVH SRXCLVDSREVTHSVHTHSISEYQRDYTGEQ LYECKGESFHISSFLFEHQRIHELDQLYSM KGCDDGFIALIPMKPRRNRAARAENPALAGSA IRCILCQGFHISSALNEHMHLHREDDLLEQS QMAEAHIPGLALTEPQRSQTERTFCAVCG ESSVNPABLADHVTVHKNEPYSFGSSYTHTS FLTEFLKGAHFYECKDCGSFISTVLTKHKE LHLEEEEBDEAAAAAAAQGEVEANVHVPQ VVLRQGLNVBAAFPSVEAAEPSVEAAEPSV EAAEPNGBAEGPDGEAAAPQRAAAQGEVEANVHVPQ VVLRQGLNVBAAFPSVEAAEPSVEAAEPSV EAAEPNGBAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGYGIEDFEBGEDGDEIQVEFYYDC HECTETTTSSTAFSEHLKTHASMIFPPANAFG ECSGYBRASTSTGGANQADEKYFKCDVCOQ LFNDHLSLARRQNTHTIG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLEREAGLSSAAVQTRIGNSA ASRSSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVAVVAVAMANNTHQ PPGSFEVPKLNVTVQDQEEHCCREGALSILQ HLRPHWDQEVTLQLFTDGTTNKLIGCYVGN TIMEDVVLNRINGNETELLYDDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFTQGEAUCHPH VCNPAFFELLANGLAKHHAHMOFWFSML WLKMGKYFSLPTGFADEDINKRFLSDFSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII VPREXQGDVGIDYSFYGYNVLANDIGNHENE FAGVSSDVDYSLYPDRELQSQWLRAYLEAYK EFKGGTEVTEKEVELFTQVNQFALASHFFW KPEVTALKVPE GUWALIQAKYSTEPTDFLGYAIVRFNQYFKM KPEVTALKVPE FFETSAKDINNVKQTPERLYDHIONSSVGK TSFLFRYADDSFTSAFVSTVGIDFEVKTYFNN EKRICLQWDTAGGGEYSTTTAYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VLIVONKCDMEDERVISTREQHLGEGLGFE FFETSAKDINNVKQTPERLYDHORDKMSESLET DPAITAAKQNTRLKSTFPPPPOPPCACC 1344 2694 A 10395 2 4136 DPRPWFWSRVDDFVTVILLISKSGHISPAKDTS						<u>'</u>	
HQRVRAGGNTSEGRSYSRSVHISLVASKPPS HNGRELVESNEGGSSITYSDLNDKRQKIPAR ENPCEGGSKNRNYEDSVIQSVPRAKPQKSVP GEGSEFKENGESSYTSDLNDKRQKIPAR ENPCEGGSKNRNYEDSVIQSVPRAKPQKSVP GEGSEFKENGGESVPSSNYEVQKARAKKK YIEHRSNETSVIHSLPFGEQTTRPRGMLYECQ ECGECFAHSSDLTEHQKHIDERKPGSRTYE WSVIRSLAPTPOTTSYAGEQYAKEQARNKCK DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE NTDGEETHISSETHQGTIEDPVQGSDMEDPQ KDDPDDKIYECEDCGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSVHTHISEQVDLYSM KGCDDGFALLPMAPRPRINARENPALAGSA IRCLLCGOGFHISSALNEHMRLHREDDLLEQS QMABEAHIPGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVIVINCHEPYEYGSSYTHTS FLTEILKGAIFFYECKDCGKSFHISTVLTKHKE LHLEEEEEDBAAAAAAAAQEVEANVHVPQ VVLRQGLNVEAAEPSVEAAEPSVEAAEPSV EAAEPNGEAEOPDGCAAEPFOEQAOPNGEAE QPNGDADEPDGVGEDPEGGEOQPYGEAE QPNGDADEPDGVGEDPEGGEOQPYGEAE GEGSYBRASTSTGGANQADEKYFKCDVCGQ LPNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPYPAPPALPEGRSQTESTSLS APAVLVVAVAVVVVVVSAVAWAMANYIHV PPGGSEVPKLNVTVQOQEEHRGGALSLQ HLRPHWDPQEVTLQLFTDGTTNKLIGCYVON TMEDVVLVRIYGNKTELLUPRGWFKSTSL) VQAHGCAPQLYCTFNNGLCYEFIQGBALDPKH VCNPARPILLARQLAKHAHANGWPKSNL WLKMGKYFSLPTGFADEDINKRFLSDFSSQI LQEEMTWMKEILSNLGSFVVLCHNDLLCKNII YNEKQDDVGTIDYESYGSTVLAAYIK EKRKLQUWDTAQGERYRTITTAYAYGAFK EKGGTGTVEREKEVELFLEVORALSHFW GUWALJQAKYSTEEPDFLGYAIVRFNQYFKM KEVEYTALKVPE GUWALJQAKYSTEEPDFLGYAIVRFNQYFKM KEVEYTALKVPE LLMTDITNEESRAVQDWSTQIKTYSWDNAQ VQLWGNKCDMEDERVISTTERQHLGEGLGFE FFETSAKDINNVKQIFTERLYDICKWSSGKG TSFLFRYADDSFTSARVSTVGIDFKVKTVFN EKRIKLQUWDTAQGERYRTITTTAYYGMAGFI LLMTDITNEESRAVQDWSTQIKTYSWDNAQ VLUVONKCDMEDERVISTTERQHLGEGLGFE FFETSAKDINNVKQIFTERLYDICKKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC	[, ,		1	
HÖĞREL VESNEK GESIYISDI.NDEK ÇÖKSP ENPCEGGSEKRNYDESUYFRAK PÇKSVP GEGSGEFK KÜDGEFSVPSSNVREYQK ARAKKK YIEHRSNETSVIHSLPFÖEQTIFAR RÖML YECQ ECOEC'A HSSOLT HENÇMEHDREK PSÖSSRNYE WSVIRSLAPT DPÖTSY AQEQY AKEQANKCK DPROFFATISED LITNIQKTYDQEK SHÖEES QË NTDGEET HSEET HIGQET TEDPVIQQES MEDEPQ KÜDPDEKTYCE COCA GEVOLT DLITDH ÇIV HOR SKICL VOSREYTHS VIHT HSISEY QROYT GEO LYECP KÇGESTHISSLE HEQDLE (1900) LYECP KÇGESTHISSLE HEQDLE (1900) LYECP KÇGESTHISSLE HEQDLE (1900) LYECP KÇGESTHISSLE HEQDLE (1900) LYECP KÇGESTHISSLE HEDDLE (1900) QWAEE ALIP (1911) LATINITA BEY QROYT GEO LYECP KÇGESTHISSLE HEQDLE (1900) QWAEE ALIP (1911) LATINITA BEY QROYT GEO LYECP KÇGESTHISSLE HEQDLE (1900) QWAEE ALIP (1911) LATINITA BEY QROYT GEO ESSVIN PALLADHIVTH KINEP YEY QEST HIT FLIEL (1911) LATINITA BE ALIP (1911) LATINITA					1		
ENPCEGGSKRRNYEDSVIQSVFRAKEÝCKSPY GEGSGGFEKDGFENYPSKEYQKARAKKK YIEHRSNETSVIRISLPFGEQTFRPRGMLYECQ ECGECFAHSSDLTEHQKIHEREKPSGRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK DFRQFFATSEDLNTNQKIYDQEKSHGESQGE NTDGETHESETHQGETIEDPVIQGSDMEDPQ KDPPDKIYECEDCGLGFVDLTDLTDHGKVH SRKCLVDSREYTHSWITHSEYQRDYTGFO LYECPKCGESFIHSSFLFEHQRIHEQDQLYSM KGCDDGFALLPMKPRRRARAERPALAGSA IRCLLCGGGFISSALNEHMRLHREDDLLEQS QMAEEAIPGLALTETPQRSQTEERIFECAVCG ESSFVIPRELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIFFYECKDCGKSFIHSTLTKHKE LHLEEGEEDBAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEFEVEAAAFEV EAAEPNGEAEGPDGEAAAPICAAGQPNGEAB QPNGDADEPDGAGIEDPEERAEPFEGKACEPE GDADEPDGVGIEDPEEGEDQE;QVEEPYYDC HECTETFTSSTAFSEHLKHASMIIFEPANNAFG ECSGYUREASTSTGGANQADEKYPKCDVCQQ LFNDH.SLARHQNTHTIG ECSGYUREASTSTGGANQADEKYPKCDVCQQ LFNDH.SLARHQNTHTIG ECSGYUREASTSTGGANQADEKYPKCDVCQQ LFNDH.SLARHQNTHTIG PGSFEVPKLINTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTIQLFTDRIKLIGCYVGN TMEDUVLVRIYGKKTELLVDROEEVKSFRVL, QAHGCAPQLVCTFNNGLCYVG UNFAIFLIARQLAKHAHHAHNGWPKSNL WLKMGKYPSLIPTGFADEDINKRITSDIPSSQI LQEEMTWMKELISNLGSVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGYSDVDYSLYPBRELQSWLRAYLEAYK EFKGFGTEVTEKEVELLFIQVNQFALASHFFW GLWALIQAKYSTIETPGFLGYBURASHFFW GLWALIQAKYSTIETPGFLGYBURASHFFW GLWALIQAKYSTIETPFLGYAURFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSDQNFDYMFKLLLIGNSSVGK TSFLRYADDSFTSAFVSTYGIDFKKTVFKN EKRIKLQIWDTAGQERYRTTITAYYRGAMGFI LMYDITNEESSNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGFLEFFETSAKDNINVKQTFERYTTITAYYRGAMGFI LMYDITNEESSNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGFLEFFETSAKDNINVKQTFERYTTITAYYRGAMGFI LMYDITNEESSNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGFLEFFETSAKDNINVKQTFERYTTITAYYRGAMGFI LMYDITNEESSNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGFLEFFETSAKDNINVKQTFERYTTITAYYRGAMGFI LMYDITNEESSNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGFLEFFETSAKDNINVKQTFERYTJUICDKMSESLET DPAITAAKQNIRLKETPPPPQPNCAC DPAPFMSRVDDPVYNLLIBLSSKGHIEPAKKDIS		1				·	
ENPCEGGSKRRNYEDSVIQSVFRAKEÝCKSPY GEGSGGFEKDGFENYPSKEYQKARAKKK YIEHRSNETSVIRISLPFGEQTFRPRGMLYECQ ECGECFAHSSDLTEHQKIHEREKPSGRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK DFRQFFATSEDLNTNQKIYDQEKSHGESQGE NTDGETHESETHQGETIEDPVIQGSDMEDPQ KDPPDKIYECEDCGLGFVDLTDLTDHGKVH SRKCLVDSREYTHSWITHSEYQRDYTGFO LYECPKCGESFIHSSFLFEHQRIHEQDQLYSM KGCDDGFALLPMKPRRRARAERPALAGSA IRCLLCGGGFISSALNEHMRLHREDDLLEQS QMAEEAIPGLALTETPQRSQTEERIFECAVCG ESSFVIPRELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIFFYECKDCGKSFIHSTLTKHKE LHLEEGEEDBAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEFEVEAAAFEV EAAEPNGEAEGPDGEAAAPICAAGQPNGEAB QPNGDADEPDGAGIEDPEERAEPFEGKACEPE GDADEPDGVGIEDPEEGEDQE;QVEEPYYDC HECTETFTSSTAFSEHLKHASMIIFEPANNAFG ECSGYUREASTSTGGANQADEKYPKCDVCQQ LFNDH.SLARHQNTHTIG ECSGYUREASTSTGGANQADEKYPKCDVCQQ LFNDH.SLARHQNTHTIG ECSGYUREASTSTGGANQADEKYPKCDVCQQ LFNDH.SLARHQNTHTIG PGSFEVPKLINTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTIQLFTDRIKLIGCYVGN TMEDUVLVRIYGKKTELLVDROEEVKSFRVL, QAHGCAPQLVCTFNNGLCYVG UNFAIFLIARQLAKHAHHAHNGWPKSNL WLKMGKYPSLIPTGFADEDINKRITSDIPSSQI LQEEMTWMKELISNLGSVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGYSDVDYSLYPBRELQSWLRAYLEAYK EFKGFGTEVTEKEVELLFIQVNQFALASHFFW GLWALIQAKYSTIETPGFLGYBURASHFFW GLWALIQAKYSTIETPGFLGYBURASHFFW GLWALIQAKYSTIETPFLGYAURFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSDQNFDYMFKLLLIGNSSVGK TSFLRYADDSFTSAFVSTYGIDFKKTVFKN EKRIKLQIWDTAGQERYRTTITAYYRGAMGFI LMYDITNEESSNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGFLEFFETSAKDNINVKQTFERYTTITAYYRGAMGFI LMYDITNEESSNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGFLEFFETSAKDNINVKQTFERYTTITAYYRGAMGFI LMYDITNEESSNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGFLEFFETSAKDNINVKQTFERYTTITAYYRGAMGFI LMYDITNEESSNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGFLEFFETSAKDNINVKQTFERYTTITAYYRGAMGFI LMYDITNEESSNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGFLEFFETSAKDNINVKQTFERYTJUICDKMSESLET DPAITAAKQNIRLKETPPPPQPNCAC DPAPFMSRVDDPVYNLLIBLSSKGHIEPAKKDIS	1		l '				HNGNELVESNEKGESSIYISDLNDKROKIPAR
GEGSGEFKKDGEFSVPSSNVREYQKAÑAKKK YIEHRSNETSVHSI.PFGCFIRPRGMI.YECQ ECGECFAHSSDLTEHQKIHDREKPSGSRNYE WSVIRSLAFIDPQTSYAQEQYAKEQANKCK DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE NTDGEETHSEETHGQETIEDPVIQGSMEDPQ KDDPDBKTVCECKOCI.GFVDLTDLTDHQKVH SRKCLVDSREYTHSVHTHSISEYQRDYTGEQ LYECPKCGESFIHSSFLFEHQRIHEQDQLYSM KGCDDGFIALLPMKPRRNAAERNPALAGSA IRCLLCGQGFHISSALNEHMRLHREDDLLEQS QMAEEAIJPGLALTEFGRSGTEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAJEFYECKDOESTHSTVLTKHKE LHLEEEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGINVEAAFEFVEAAFEVEAAAEPEV EAAEPNGEAAGFDGEAAAPEVEAAEPEV EAAEPNGEAAGFDGEAAAGPFGCKAEPE GDADEPDGVGIEDFEEGGEAQPPNGEAA QPNGDADEPDGAGIEDPERCAGPRGEA QPNGDADEPDGAGIEDPERCAGPRGEA ECSGYIERASTSTGGANQADEKYFKCDVCOQ HECTETFTSSTAFSEHLKTHASMIFPPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCOQ LFNDHLSLARHONTHITG 342 2692 A 10393 2 1350 GRPRSSSSDNRNFLERRAGLSSAAVQTRIGNSA ASRSFAAARPPVAPPALPRGRGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYHPY PPGSFEYFKLINTVQDQEEHRCREGALSLLQ HLRPHWDPQEVILQHTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCVEFIQGEALDPKH VCNPAIFRILARQLAKHAIRIAHNGWFKSNL WKMGKYSLIPTGRADGNKRFLSDPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIETDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QUANYGGKDSDONFDYMFKLLIGNSSVGK TSSLFRYADDSSTSAFVSTVGIDFKVTVFKN EKRIKLQIUMUTAGGERYRTITTAYYRGAMGFI LMYDITNEESSNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQOLGFE FFETSAKDNINVKQTFERLYDLICDKMSESLET DPAITAAKQNIRLKETPPPPQPNCAC DPAFFGRAMDTHLSSKGHIEPAKDTS				1] '		
YIEHRSNETSVIRSLPFGEQTFRRRGMLYECQ ECGECHARSDLTEHQKHDREKPSGSRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK DFRQFFATSEDLNTNQKIYDQKSHGEESQGE NTDGEETHSEETHGETEIDPVIQGSDMEDPQ KDPDDKIYECEDCOLGFVDLTDLTDHQKVPI SRKCLVDSREYTHSVIRTHEDPVIQGSDMEDPQ KDPDDKIYECEDCOLGFVDLTDLTDHQKVPI SRKCLVDSREYTHSVIRTHESEYQRDYTGEQ LYECPKCGESFIHSSELFEHQRHEQDQLYSM KGCDDGFIALDPMKPRRAFRAPALAGSA IRCLLCQQFHSSALNEHMRLHREDDLLQS QMAEEAIIPGLALTEFQRSQTEERLFCAVCG ESFVNPALLADHTVTHKNEPYEYGSSYTHTS FLTEPLKGAJFFYECKDCGKSFIHSTVLTKHKE LHLEEEEEDBAAAAAAAAQEVEANVHVPQ VVLRIGGLNVEAAFEVVAAAPEVGAAEPEV EAAFPNGEAAGPDGEAAAPEVAAAQEVEANVHVPQ VVLRIGGLNVEAAFEVEAAEPEVGAAEPEV GDADEPDGVGIEDPEGGEDQEIQVEPYYDC HECTETTTSSTATSEHLKTHASMIIFEPANAFG ECSGYJERASTSTGGANQADEKYFKCDVCQQ LFDHHSLARHQNTHITG ECSGYJERASTSTGGANQADEKYFKCDVCQQ LFDHHSLARHQNTHITG I342 2692 A 10393 2 1350 GRPFSSSDNRIFLERAGLSSAAVQTIGIONSA ASRRSPAARPPVPAPPALPRGFGTEGSTSLS APAULVVAVAVVVVVVSAVAWAMANYTHV PGSFEVFKLINVTVQDQEEHRCREGALSLLQ HLRPHWPDGVVLVRIYGNKTELLVDRDEEKYSFRVL QAHGCAPQLVCTFNNGLCYVGN TMEDVVLVRIYGNKTELLVDRDEEKYSFRVL QAHGCAPQLVCTFNNGLCYFGQGEALDPKH VCNPAFFRLIARQLAKHABHAHNGWIFKSNL WLKMGKYFSLIPTGFADEDDINKFISDIPSSQI LQEMTWMKEILSNLGSFVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNIFINE FAGVSDVDYSLYPDRELQCVRFAVLEAYK EFKGFGTETTEKEVEILFJQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE FRAGSDVDVSLYPDRELQCWLRAYLEAYK EFKGFGTETTEKEVEILFJQVNGFALASHFW GLWALIQAKYSTIETDFLGYNGFALASHFW GLWALIQAKYSTIETDFLGYNGFALASHFW GLWALIQAKYSTIETDFLGYNGFALASHFW KPEVTALKVPE EKRIKLQIWDTAGQERYRATITTAYYYGAMGFI LMYDITNEESRNAVQDWSTQIKTYSWDMAQ VILVGNKCDMEDERVISTERGQOHLGFQLGFE FFETSAKDNINVKQTFERLVTILGNSSVGK TSFLFRYADDSFTSAFVSTYGIDFKKTVTVFKN EKRIKLQIWDTAGQERYRTITTAYYYGAMGFI LMYDITNEESRNAVQDWSTQIKTYSWDMAQ VILVGNKCDMEDERVISTERGQOHLGFQLGFE FFETSAKDNINVKQTFERLVTDLCKAKGTIST							
ECGECFAHSSDLITHQKHIPDRERPGSRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE NTDGETHSEETHGQETIEDPVIQGSDMEDIPQ KDPPDDKIYECEDCGLOGFOLIDLITHDQKYH SRKCLVDSREYTHSVHTHISISEVQRDYTGEQ LYECPKGGSFHISSTLFHQRHEDQDLYSM KGCDGFIALI PMKPRRNRAAERNPALAGSA RCLLCQQFHISSALNEHMRLHREDDLLEQS QMAEEAIIFGLALTEFQRSQTEERLIFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLIFELKGAIFFYECKCGKSFHISTVLTKHKE LHLEEEEDBAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEPEGKAEEPE GDADEPDGVGIEDPEGEDQDEIQVEEPYYDC HECTETTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGAAQAEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYIHV PGSPEVPKLNVTVQOEERCREGALSLQ HLRPHWDPQSVTLQLFTDGITNKLIGCYVGN TMEDVVLLVRIYGNKTELLVDRGERKSGESLSLQ HLRPHWDPQSEVTLQLFTDGITNKLIGCYVGN TMEDVVLLVRIYGNKTELLVDRGERKSGESLSLQ LQEEMTWMKELISNLGSPVVLCHDDLLCKNII VNEKQGDVQFIDYESGYNYLAYDIGNHFNE FAGVSDVDYSLYPDELQSQWLRAYLEAYK EFKGFGTEVTEKEVELLFIQNVQFLASHFFW GLWALIQAKYSTEPTPFLQYAIVERNQYFKM KPEVTALKVPE FEGSFGTEVTEKEVELLFIQNVQFLASHFFW GLWALIQAKYSTEPTPFLGYAIVERNQYFKM KPEVTALKVPE FEGSFGTEVTEKEVELLFIQNVQFLASHFFW GLWALIQAKYSTEPTPFLGYAIVERNQYFKM KPEVTALKVPE FFETSAKDNINVKQTFERLVDILCIGKNSSCIK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN KRIKLQUIDTAGQBRYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNCOMEDBERNSTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDILCIGKMSSLET DPAITAAKQNITRLETPPPQCNACC DPAFMSRSVDDFYNTLHILSSKGHISPAKDTS			[1	·	VIER SMETSVILLGI DECEUTEDBD CAR VECO
WSVIRSLAPIDPQTSY'AQEQYAKEQARNIKCK DFROFFATSEDLNTNQKIYDQEKSHGESQGE NTDGETHSEETHGQETTEDPVIQGSSMEDPQ KDDPDDKIYECEDCGLGFVDLTDLTDHQWY SRKCLVDSREYTHSVHTHSISEYQRDYTGEQ LYECPKCGESFHISSFLFEHQRHEGDQLYSM KGCDDGFIALLPMKPRRNRAAERNPALAGSA IRCLLCQQGFHISSALNEHMRLHREDDLLEQS QMAEEAIIRGLALTEFQRSQITEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFHISTVLTKHKE LHLEEEEEDEAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPFVEAAEPEV EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGGADEPDGAGIEDERAEEPEGKAEEPE GDADEPDGVGIEDPEEGEDQEIQVEPYYDC HECTEFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCCQ LFNDHLSLARHQNTHTG GRPRSSDNNPKIFEREAGLSSAAVQTRIGNSA ASRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVVVVVVVAVAWAMANYIHV PPGSPEVPKLNYTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNSLIGCYYGN TMEDVVLNIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHHAHAHNGWIPKSNL WLKMGKYFSLPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSFVVLCHNDLLCKNII YNEKGGDVOFIDVEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELOSQWLRAYLEAYK EFKGFGTEVTEKEVELLFIQVNQFALASHFFW GLWALIQAKYSTEEPDFLGYATVRRNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTIMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIGNSSVGK TSFLIRYADDSFTSAFVSTVGIDFKKTTVFKN KERKILQUIDDTAGGERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQKKTYSWDNAQ VILVGNCOMEDEERVSTTERGQHLGEOLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITRLKETPPPPQNACC 1344 2694 A 10395 2 4136 DRPWNSRVDDFVINLHLISSKGHISPAKDTS			l	· '		 	
DFRQFFATSEDLINIQKIYDQEKSHGESQGE NTDGEFHSSETHGQETISDEPLQGSMDEDQ KDDPDDKIYECEDCGLGFVDLTDLTDHQKVH SRKCLVDSREYTHSVHTHSISEVQRDYTGEG LYECFKCGSEFHISSET,FEHQRIHEQDQLYSM KGCDDGFIALLPMKPRRINARAENPALAGSA RCLLCGGGFHISSALHMI.HREDDLLEQS QMAERAIIPGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVIVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFHISTVLTKHKE LHLEEEEDEAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAPEPVEAAAFPEVEAAEPEVE EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEGEGDQELQVEEPYYDC HECTETTSSTAFSEHLKTHASMIFIEPAANAFG ECSGYIERASTSTGGAAQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYIHV PPGSPEVPKINVTVQDEFIRCKEGALSLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYYGN TMEDVVLVRIVGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNIGLCYEFIQGEALDFKH VCNPAIFRLIARQLAKHAHHAHNGWIPKSNL WLKMGKYFSLPTOFADEDINKRYLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKCGDVOFIDVEYSDYNLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELLFIQNVQFALASHFFW GLWALIQAKYSTEPDFLGYATVRNQYFKM KPEVTALKVFE GLWALIQAKYSTEPDFLGYATVRNQYFKM KPEVTALKVFE LMYDITNEESFNAVGTWSTVSTVGIDFKVKTVFKN KPEVTALKVFE EKRIFLQUIDTAGGREYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNCOMEDDERYSTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNCOMEDDERYSTGRGQHLGFGLGFE FFETSAKDNINVKQTFERLVDILICDKMSESLET DPAITAAKQNITRLKETPPPPQCNACC DPFWNSRSVDDFYNTLHJLISSKGHIEFBAKDTS	1 1		ĺ	ĺ	1	ľ	
NTDĞEETHSEETHGQËTIEDPÜQGSDMEDPQ KDDPDKIYECEDCGLGFVDLTDLITDHQKVH SRKCLVDSREYTHSVHTHSISEYQRDYTGEQ LYECPKCGESFIHSSVHTHSISEYQRDYTGEQ LYECPKCGESFIHSSVHTHSISEYQRDYTGEQ LYECPKCGESFIHSSTLFHQRIHEQDQLYSM KGCDDGFIALLPMERNRAAGERPALAGSA IRCLLCGGGFHSSALNEHMRLHREDDLLEQS QMAEEAIIPGLALTEFQRSQTEERLFECAVCG ESFVNPABLADHVTHKNEPYEVGSSYTHTS FLITEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEBEDEAAAAAAAAQEVEAAPISVEAAPEPV EAAEPNGEAEGPDGEAAEPIGEAQQPNGEAE QPNGBADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEERGEDQEQVEEPYYDC HECTETTFSSTATSEHLKTHASMIFPEANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARIQNITHIG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARIQNITHIG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARIQNITHIG PGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRHWDPQGVTLQLFTDGITMKLIGCVYGN TMEDVVLVRYGNKTELLVDRDEEVSFRVL. QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAHFILIRQLAKHAHAHHANGWFKSNL. WLKMGKYFSLITGFADEDINKRFLSDIPSSQI LQEEMTWMKELSNLGSPVVLCHNDLLCKNII YYEKQGDVQFIDYSTGYNVLAYDIGNHFNE FAGVSDVDVSLYPDRELQSOWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTEIPGFADEDINKRFLSDIPSSQI LQEMTWMKELSNLGSPVVLCHNDLLCKNII YYEKQGDVQFIDYSTGYNVLAYDIGNHFNE FAGVSDVDVSLYPDRELQSOWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTEIPGTAGTAYHRNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGDFKVKTVFKN EKKIKLQIWDTAGGERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTIGTTYSWDNAQ VILVGNKCDMEDBERVISTERGOHLGFOLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPRQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPQNCAC DPAITAAKQNTRLKETPPPPRONCAC DPAITAAKQNTRLKETPPPPRONCAC DPAITAAKQNTRLKETPPPRONCAC DPAITAAKQNTRLKETPPPRONCAC DPAITAAKQNTRLKETPPPRONCAC DPAITAAKGRITES							
SRKCLVDSREYTHSVIRTHSISEYQRDYTGEQ				[1	[
SRKCL/VDSREYTHSVHTHSISEYQRDYTGEQ LYECPKCGESFHSSFLFEHQRIHEQDQLYSM KGCDDGFIALLPMKPRRNRAAERNPALAGSA IRCLLCQGGFHSSALNEHMRLHREDDLLEQS QMAEAHIPGLAIT-PGRSQTEERLEPCAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLIEPLKGAJPFYECKDCGKSFHSTVLTKHKE LHLEEEBEDEAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPTOGAGQPNGEAE QPNGDADEPDGAGIEDPERFAEEPEGKAEEPE GDADEPDGVGIEDPERGEDEPEKAEEPE GDADEPDGVGIEDPERGEPEGKAEEPE GDADEPDGVGIEDPERGAEEPE GDADEPDGVGIEDPERGAEEPE GDADEPDGVGIEDPERGAEEPE GDADEPDGVGIEDPERGAEEPE GDADEPDGVGIEDPERGAEPEPE GDADEPDGVGIEDPERGAEEPE GDADEPDGVGIEDPERGAEPEP GDADEPDGVGIEDPERGAEPEP GDADEPDGVGIEDPERGAEPEP GDADEPDGVGIEDPERGAEPEP GDADEPDGVGIEDPERGAEPEP GDADEPDGVGIEDPERGAEPEP GDADEPDGVGIEDPERGAEPERGAE ECSGVIERASTSTGGANQADEKYFKCDVCQ LFNDHLSLARHQNTHTG ECSGVIERASTSTGGANQADEKYFKCDVCQ LFNDHLSLARHQNTHTG PGSFEVPKLNVTVQDQEEHRCREGALSLQ HRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVAVAVVVVVVAVAWAMANNYIHV PPGSFEVPKLNVTVQDQEEHRCREGALSLQ HAPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTBNGLCYEFIQGEALDPKH VCNPAFFELLARQLAKHAHAHANGWIPKSNL WLKMGKYFSLIPTGG ADEDINKRFLSDIPSSQI LQEEMTWMKELISNLGSPVVLCHNDLLCKNII YNEKQGDVQPIDYEYSGYNVLAYDIGNHFNE PAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELLFIQVNOPALASHFFW GLWALQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE GLWALQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE GLWALQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE GLWALQAKYSTIEFDFLGYAIVRFNQMASA QDARYGQKDSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGUDFKVKTVFKN EKRIKLQIWDTAGQERYRITITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGGHLGEQLGFE FFETSAKDNINVKQTFERLVDICDKMSESLET DPAITAAKQNITLKETPPPPPONCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFYNNLHLISSKGHISPAKDTS				[}		NTDGEETHSEETHGQETIEDPVIQGSDMEDPQ
SRKCL/VDSREYTHSVHTHSISEYQRDYTGEQ LYECPKCGESFHSSFLFEHQRIHEQDQLYSM KGCDDGFIALLPMKPRRNRAAERNPALAGSA IRCLLCQGGFHSSALNEHMRLHREDDLLEQS QMAEAHIPGLAIT-PGRSQTEERLEPCAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLIEPLKGAJPFYECKDCGKSFHSTVLTKHKE LHLEEEBEDEAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPTOGAGQPNGEAE QPNGDADEPDGAGIEDPERFAEEPEGKAEEPE GDADEPDGVGIEDPERGEDEPEKAEEPE GDADEPDGVGIEDPERGEPEGKAEEPE GDADEPDGVGIEDPERGAEEPE GDADEPDGVGIEDPERGAEEPE GDADEPDGVGIEDPERGAEEPE GDADEPDGVGIEDPERGAEEPE GDADEPDGVGIEDPERGAEPEPE GDADEPDGVGIEDPERGAEPEPE GDADEPDGVGIEDPERGAEPEPE GDADEPDGVGIEDPERGAEPEP GDADEPDGVGIEDPERGAEPEP GDADEPDGVGIEDPERGAEPEPE GDADEPDGVGIEDPERGAEPEP GDADEPDGVGIEDPERGAEPEP GDADEPDGVGIEDPERGAEPERA PECTETTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCQ LFNDHLSLARHQNTHTG PGSEVPKLNVTVQAVAVVAVVAVAWAMANNYIHV PFGSFEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVAVAVVVVVVAVAWAMANNYIHV PFGSFEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTBNGLCYEFIQGEALDPKH VCNPAFFELLARQLAKHAHAHNGWIPKSNL WLKMGKYFSLIPTGG ADEDINKRFLSDIPSSQI LQEEMTWMKELISNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNVLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELLFIQVNOFALASHFFW GLWALQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE GLWALQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE GLWALQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE GLWALQAKYSTIEFDFLGYAIVRFNQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRITITTAYYRGAMGFI LMYDITNEESINAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGGHLGEQLGFE FFETSAKDNINVKQTFERLVDICDKMSESLET DPAITAAKQNITLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFYNNLIHLISSKGHISPAKDTS				1	!	!	KDDPDDKIYECEDCGLGFVDLTDLTDHOKVH
LYECPKC GESFHASFLFEHQRIHEQ DQL YSM KGCDDGFIALL PMKPRRNRA AERNPALAGSA IRCLLCQGFHISSALNEHMRLHREDDLLEQS QMAEEAIIPGLALTEFQRSQTEERLFECAVCG ESFNPAELADHVHVPVIVNKNEPPYSGSYTHTS FLTEPLKGAIPFYECKDCGKSFHSTVLTKHKE LHLEEEEBDEAAAAAAAAQEVEANHVHVQ VVLRIQGLINVEAAEPEVEAAEPEVEAEPEVE EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGGAGEDPEERAEEPEGKAEPEP GDADEPDGVGGEPPEGEDQEIQVEEPYYDC HECTETTTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTAFSEHLKTHASMIIFEPANAFG ASRRSSASRPAPPPVIQLCEFINGLCAEFICGGALDPKH VCNPAIFRLIARQLAKHAHAHANGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEHLSNLGSPVULCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVISDVOYSL YPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELIFIQVNQFALASHIFFW GLWALIQAKYSTEEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLIPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKWTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVONKCDMEDERVISTERGQHLGFDLGFE FFETSAKDNINVKQTFERI VDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLISKSGHISFAKDTS				. '	1	1	SRKCLVDSREYTHSVIHTHSISEYORDYTGEO
KGCDDGFIALLPMKPRRNRÅAERNPÄLAGSA IRCLLCQQGFIHSSALNEHMRLHREDDLLEQS QMAEEAIIPGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEDEAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAEPEGKAEEPE GDADEPDGGEDPEEGEGQEIQVEEPYYDC HECTETTSSTAFSEHLKTHASMIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCQQ LFNDHLSLAHQNTHTG ECSGYIERASTSTGGANQADEKYFKCDVCQQ LFNDHLSLAHQNTHTG ECSGYIERASTSTGGANQADEKYFKCDVCQQ LFNDHLSLAHQNTHTG PRESSSDNRNFLERERGGLSSAAVQTRIGNSA ASRRSPAARPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVWAMANYIHV PRESSPEVPKLINVTVQDQEEHECREGALSLQ HLRPHWDPQEVTLQLFTDGTINKLIGCVVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCVEFIQGEALDPKH VCNPAIFRLIARQLAKIHAHAHNGWIPKSNL WLKMGKYFSLPTGFADEDINKRFLSDIPSSQI LQEEMTWMKELISNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAVDIONHFNE FAGVSDVDYSLYPDRELQSQWRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE EFFTSAKDNIFVSTVGIPGKWTVYKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVONKCOMEDRENTISTERGQMLGGELGFE FFETSAKDNIFVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC DPAPTMAKQNTRLKETPPPPPPQPNCAC DPAPTWNSRVDDFVTNLIHLISSKGHISFAKDTS	l I			i i		1	
IRCLLCGQGFIHSSALNEHMRLHREDDLLEQS QMAEAIIPGLALTEFQRSQTEERLFECAVCG ESFVNPAELADHVIVHKNEPYEYGSSYTHITS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE LHLEEEEDEAAAAAAAQEVEANVHVPQ VVLRIQGINVEAAEPEVEAAEPEVEAAEPEV EAAEPIGEAEGPDGEAEPEGKAEEPE GDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGAGIEDPEGEDGEIQVEEPYYDC HECTEITTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCOQ LFNDHLSLARHQNTHTG GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAMAMNYIHV PPGSFEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFIDGITNKLIGCYVGN TMEDVVLVIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCVEFIQGEALDFKH VCNPAIFRLIARQLAKHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKWKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDRVISTERQQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICCKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC	1						
QMAEEAIĪPGLALTEFQRSQTERLFECAVCĞ ESFVNPAELADHVIVHKNEPYEYGSSYTHTS FLTEPLKGALTPFYCKDCGKSFIHSTVLTKHKE LHLEEBEEDEAAAAAAAQEVEANVHVPQ VVLRIQGINVEAAEPEVEAAEPEVEAEPEVE EAAEPNGEAEQPDGEAAEPIGEAGQPNOEAE QPNGDADEPDGVGIEDPEEGEDQEIQVEEPYYDC HECTEITISTATSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVAPPAPLPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMNYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTIQLFTDGTINKLIGCYVGN TMEDVVLVRYGNKTELLVDRDEFVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHAHHAHNGWIPKSNL WLKMGKYPSLDTGFADEDINKRFLSDPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRRNQYFKM KEPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIGNSSVGK TSSLFRYADDSFTSAFVSTVGDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDILCDKMSESLET DPAITAAKQNTRLKETPPPPPPCNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVYNLHLISSKGHISPSAKDTS			ł	1	ľ	}	
ESFVIPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFHSTVLTKHKE LHLEEEEDEBAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAEPEVEAAEPEVEAEPEVEAEPEV EAAEPNGEAEGPDGEAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEEGEDDEIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYPKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSDNRNFLRERAGLSSAAVQTRIGNSA ASRSPAARPVPAPPALPRGRPGTEGSTSLS APAVLVAVAVVVVVSAVAWAMNYHVV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTILQLFTIGGITNKLIGGVYGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFELIARQLAKHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVOPJYSYGSYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELFIQVNQFALASHFFW GLWALIQAKYSTEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDDITNEESFNAVQD WSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNNVKQTFERL VDIICDKMSESLET DPAITAAKQNTRLKETPPPPPPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLHLISSKGHISPAKDTS			1	1		1	
FLTEPLKGAIPFYECKDCGKSFHISTVLTKHKE LHLEEEEDEAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAAPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEGEEDQEIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVAWAMANYIHV PPPGSPEVPKLNVTVQDQEEHRCREGALSLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGINKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYPSLPTGFADEDINKRFLSDIPSSQI LQEENTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNVLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			İ		!	,	QMAEEAIIPGLALTEFQRSQTEERLFECAVCG
LHLEEEEDEAAAAAAAAAQEVEANVHVPQ VVLRIQGLNVEAAAEPEVEAAEPEVEAAEPEV EAAEPRIGEAGPDOEAAEPEVEAAEPEVEAAEPEV EAAEPRIGEAGPDOEAAEPIQEAGPORIGEAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGAGIEDPEEGEDQEIQVEEPYYDC HECTETFISSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNITHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRFILREAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFFALIARQLAKIHAIHAHNGWPKSNL WLKMGKYFSLPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQEDVQFIDVEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHIFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGGKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGGHLGEQLGFE FFETSAKDNINVKQTFERLVDICDKMSESLET DPAITAAKQNITRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS				1	1		ESFVNPAELADHVTVHKNEPYEYGSSYTHTS
VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAGPDGEAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGAGIEDPEEGEDQEIQVEEPYYDC HECTETITISTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHIAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFFKGFGTEVTEKEVEILFIQVNOPALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPPPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLHILSSKGHISPAKDTS	1		ì	1	{	1	FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE
VVLRIQGLNVEAAEPEVEAAEPEVEAAEPEVEAAEPEVEAAEPEV EAAEPNGEAEGPDGEAGPDGEAE QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGAGIEDPEEGEDQEIQVEEPYYDC HECTETITISTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKHIAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFFKGFGTEVTEKEVEILFIQVNOPALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPPPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLHILSSKGHISPAKDTS			ł	1	}		LHLEEEEEDEAAAAAAAAAOEVEANVHVPO
EAAEPNGEAEGPDGE.AAEPIGEAGQPNGEAE QPNGDADEPDGAGIEDPERAEPPEGKAEEPE GDADEPDGVGIEDPEGGEDQEIQVEEPYYDC HECTETFTSSTAFSEHLKTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG GCSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG GRPRSSSDNRNFLEERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVVSAVWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCVEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRPLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNESSFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPPPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLHILSSKGHISPAKDTS	ł		ł		}		
QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC HECTETTSTSTAFSEHI.KTHASMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDRNFIRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLINVTVQDQEEHRCREGALSLLQ HI.RPHWDPQEVILQETDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFYTNLIHLSSKGHISFAKDTS	i ·					ľ	
GDADEPDGVGIEDPEGEDQEIQVEEPYYDC HECTETITSSTAFSEHLKTHASMIIFERANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG 1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNBKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KEPVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDTINEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPPQNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS				,	i l		OBJECT TO EARCH DO EARCH A PERFECT A PERFE
HECTETFTSSTAFSEHLKTHÄSMIIFEPANAFG ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNITHTG GRPSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TIMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAHHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPPOPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLHILSSKGHISPAKDTS						,	
ECSGYIERASTSTGGANQADEKYFKCDVCGQ LFNDHLSLARHQNTHTG GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLL_LFTDGITINKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYPSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVELLFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDILCDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLHLSSKGHISFAKDTS	1				1	1	
1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGROTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTA YYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDER VISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC				1		!	
1342 2692 A 10393 2 1350 GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAMAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFITDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYPSLPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHIFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	ł		1			ļ	ECSGYIERASTSTGGANQADEKYFKCDVCGQ
ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1			1			LFNDHLSLARHQNTHTG
ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVILQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1342	2692	Α	10393	2	1350	GRPRSSSDNRNFLRERAGLSSAAVOTRIGNSA
APAVLVVAVAVVVVVSAVAWAMANYIHV PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HLRPHWDPQEVILQLFTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			l '		1		
PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ HI.RPHWDPQEVTI.QI.FTDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQEGALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFI.SDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHIFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS				i . '	1		APAVI VVAVAVVVVVVVVA AVAMAAANOTIN
HLRPHWDPQEVILQLFIDGITNKLIGCYVGN TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHIFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1		1				
TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1		ł	1			
QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			1				
VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1						
VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS					'	. !	QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH
WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1					1	VCNPAIFRLIAROLAKIHAIHAHNGWIPKSNI.
LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1			[[1	WLKMGKYFSLIPTGFADEDINKRFI SDIPSSOI
YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	ł		1	}			
FAGVSDVDYSLYPDRELQSQWLRAYLEAYK EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			1				
EFKGFGTEVTEKEVEILFIQVNQFALASHFFW GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1 .						
GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNITLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	·						
KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS				j i	1		
KPEVTALKVPE 1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS							GLWALIQAKYSTIEFDFLGYAIVRFNOYFKM
1343 2693 A 10394 102 839 PEAQTSAVLAREKGHLPTMRHEAPMQMASA QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS						l	KPEVTALKVPE
QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1343	2693	A	10394	102	839	
TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1		١.,	'''''			
EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			1				
LMYDITNEESFNAVQDWSTQIKTYSWDNAQ VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			l		1		
VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1					1	EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI
VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	1		1				LMYDITNEESFNAVQDWSTQIKTYSWDNAO
FFETSAKDNINVKQTFERLVDIICDKMSESLET DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			1				
DPAITAAKQNTRLKETPPPPQPNCAC 1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS	,]]	j '	J J	ļ	
1344 2694 A 10395 2 4136 DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS			1	f !	1		
	137	2604	<u> </u>	10205	 	4126	
LOORTPAEMSPVLHFYVRPSGHEGAASGHTR							
	1344	2094	Λ.	10373	۱ <i>-</i> ۱	7130	

residue of peptide sequence peptide sequence peptide sequence sequ	SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid	Predicted end nucleotide location corresponding to last amino acid residue of peptide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
RKLOGKLPELOGVETELCYNWWTAEALPSA EETKKLMWLGGCHLDDVARSBWLIGSN DLLLEVOPRINSTPTSTNIVSVCRATGLGPV DRWETTERYRLSFAHPPSAEVEAIALATLEDR MTEGHFPFPIGSTSESMPEPLNGFINLIGEGR LALBEKANGELGLALDSWULDFYTKFOELOR NPSTVEAFDLAGSNSEHSRHWFFKGCLHVDG QKL VISISJESSIMSTOESSNIPNNIVLEFOENSA IQGEVEFLRFEDFTEPSRFQQQGGLRIVVFT AETHNFPTGVCPSGGSTTOTIGGRIRDVCCTG RGAHVVAGTAGYCFGNLHIEGYNLPWEDLSF QYPGHFAPPLEVALBSANGASDYGNIFGEPV LAGFARSLGLJPDGGREEWIKPIMFSGGIGS MEADHISKEAPPEGMEVVXUGGPYYRIGVGG GAASSVQVQGONTSDLDFGAVORGDFBMEQ KMMRVYRACVEAPKOMPCSHDOGAGMO NNIVLELSDPAGAITYTSFFQLGDFTLANAEIN GAEYQESNALLLESPRDETLTHVSAERCPA CEVGTTTGDRRIVLYDDRECEVERNOGGOAP PTPFTPYDLELEWYLGKMPKEFFLGVKPP MLOPLALPOLSVFGALERVALPAVASKRY LTDRVDRSVGGLVAOQCOVEPLOTI-ADVA VALSHEELIGAATALGEOPVKSLDPKVAA RAVAGAELTRILVFALVTDLRVKCSGNWM WAAKLPGGGAALADACEAMVAWMAALGVA VAGAGESISMAARVGTTEVTAPOSL VISAYA VCPDITATYTPDLRHEGRGRILLYVALSPGQ HELGGTALAGCTSQLGHPPDLDLPBNLVRA FSITQGLLKDBLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVFVVFVOVSLVLFAEEPCLV LEVQEPDLAGVLKRYRDAGLIKCLEGFTGE AGFRAWRYSNGGLAVOVTAQDLCSGNAGULT FREGVAFYGGGSYADVLGSAKGWAAAVTFL RGVAFFGGLARDAGARVLSGMAGAUCHT RGVAFYGGGSYADVLGSAKGWAAAVTFL RGVAFFGGLARDAGAGAUCHT RGVAFYGGGSYADVLGSAKGWAAAVTFL RGVAFFGGGAALADACEAMVROWAGAUCHT RGVAFYGGGSYADVLGSAKGWAAAVTFL RGVAFYGGGSYADVLGSAKGWAAAVTFL RGVAFFGGAALGADACEAMVROWAGAUCHT RGVAFYGGGSYADVLGSAKGWAAAVTFL RGVAFYGGGSYADVLGSAKGWAAAVTFL RGVAFFGGAALGADACEAMVROWAGAUCHT RGVAFYGGGSYADVLGSAKGWAAAVTFL RGVAFFGGAALGADACEAMVROWAGAUCHT RGVAFYGGGSYADVLGSAKGWAAAVTFL RGVAFFGGAALGADACEAMVROWAGAUCHT RGVAFYGGGAALGADCGACHT RGVAFYGGGSYADVLGSAKGWAAAVTFL RGVAFYGGGTAAACKGGSGCACHT RGVAFYGGGSYADVLGSAKGWAAAVTFL RGVAFFGGAALGADCGAUCHT RGVAFYGGGAACHT RGVAFYGGGGAACHT RGVAFYGGGAACHT RGC					peptide	sequence	/=possible nucleotide deletion, \=possible
1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDGSDFQHRE RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVSI MYLVLVLAVQVHAWQLYYSKKLLDSWFTST QEKKHK 1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRGT EIGVLAKAFIDQGKLIPDDVMTRLALHELKNL TQYSWLLDGFPRTLPQAEALDRAYQIDTVINL NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP 1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEL							RKLQGKLPELQGVETELCYNVNWTAEALPSA EETKKLMWLFGCPLLLDDVARESWLLPGSN DLLLEVGPRLNFSTPTSTNIVSVCRATGLGPV DRVETTRRYRLSFAHPPSAEVEAIALATLHDR MTEQHFPHPIQSFSPESMPEPLNGPINILGEGR LALEKANQELGLALDSWDLDFYTKRFQELQR NPSTVEAFDLAQSNSEHSRHWFFKGQLHVDG QKLVHSLFESIMSTQESSNPNNVLKFCDNSSA IQGKEVRFLRPEDPTRPSRFQQQQGLRHVVFT AETHNFPTGVCPFSGATTGTGGRIRDVQCTG RGAHVVAGTAGYCFGNLHIPGYNLPWEDLSF QYPGNFARPLEVAIEASNGASDYGNKFGEPV LAGFARSLGLQLPDGQRREWIKPIMFSGGIGS MEADHISKEAPEPGMEVVKVGGPVYRIGVGG GAASSVQVQGDNTSDLDFGAVQRGDPEMEQ KMNRVIRACVEAPKGNPICSLHDQGAGGNG NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW GAEYQESNALLLRSPNRDFLTHVSARECPA CFVGTITGDRRIVLVDDRECPVRRNGQGDAP PTPPPTPVDLELEWVLGKMPRKEFFLQRKPP MLQPLALPPGLSVHQALERVLRLPAVASKRY LTNKVDRSVGGLVAQQQCVGPLQTPLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGVA VDGGKDSLSMAARVGTETVRAPGSLVISAYA VCPDITATVTPDLKHPEGRGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWEE TSFQLDRLQAEPRCVAEEERGLRERMGPSYC LPPTFPKASVPREPGGPSPRVAILREGSNGDR EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFHP RAGAELRRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLLRHNL SGRYESRWASVRVGPGPALMLRGMEGAVLP VWSAHGEGYVAFSSPELQAQIEARGLAPLHW ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPFFDTLTTS
1346 2696 A 10398 1 718 DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRGT EIGVLAKAFIDQGKLIPDDVMTRLALHELKNL TQYSWLLDGFPRTLPQAEALDRAYQIDTVINL NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP 1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEL	1345	2695	A	10396	65	642	RVAMHYQMSVTLKYEIKKLIYVHLVTWLLLV AKMSVGHLRLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVSI MYLVLVLAVQVHAWQLYYSKKLLDSWFTST
QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP 1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLIENTNGEL	1346	2696	A	10398	1	718	DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRGT EIGVLAKAFIDQGKLIPDDVMTRLALHELKNL TQYSWLLDGFPRTLPQAEALDRAYQIDTVINL NVPFEVIKQRLTARWIHPASGRVYNIEFNPPK
I I VANPEAUNICATION V V V V V V V V V V V V V V V V V V V	1347	2697	A	10402	153	1969	QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP

NO: of No: of nucle eotide sequence unce unce unce unce unce unce unce	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence USSN Ogada Ogresponding Infection Sequence Sequence Ogada O	NO: of	NO: of	hod	ID NO:		nucleotide	D=Aspartic Acid, E=Glutamic Acid,
uence wence wence where the service of the service of the service of peptide amino acid residue of peptide sequence where sequence were sequence where sequence where sequence were sequence where sequence where sequence were sequence where sequences are sequence where sequences where sequence where sequences where sequences are sequenced where sequences where sequences where sequences are sequences and sequence where sequences where sequences are sequences and sequence sequences where sequences are sequences and sequences where sequences are sequences as sequence sequences and sequences are sequences and sequences are sequences and sequences are sequences as sequences and sequences are se			İ				
uence 914 a	1			1			
amino acid residue of peptide sequence		uence					
residue of peptide peptide peptide peptide peptide peptide peptide peptide sequence Y=Tyrosine, X=Unknown, **=Siop codon, pepsible uncleotide insertion LMNLAGKNRGFSLGSTVKSHTKGIWMWCV	uence	1	ĺ	914	, –		
peptide sequence mucleotide deletion, "possible mucleotide insertion LMNRLAGKINGFSLGSTVKSHTKGIWMWCV PHPKKPEHTLVLLDTEGLGDVKKGONONDS WIFITAVLLSSTLVYNSMGTINQQAMDQLYY VTELTHRIRSKSSPDENEDSADFVSFFPDFY WTI.RDFSLDLEADGQPLTPDEYLEYSLKLTQ GTSQKDKNFNLPRLCTRKFFFKKKCFVPDLPI HRKALAQLEKLQDEELDPEFVQQVADFCSYI FSNSTKTLSGGIKVNGPRJESLVLTYINAISR GDLPCMENAVLALAQIENSAAVQKAIAHYD QQMGGKVQLPAFTLQELDHRYSEREATEV YMKNSFKDVDHLFQSKRTAAHYD QQMGGKVQLPAFTLQELDHRYSEREATEV YMKNSFKDVDHLFQSKRTAAGTSK PGGYCLFfQKLQDLEKKYYEPRKGIQAEELL QTYLKSKESVTDAILQTDQILTEKKEEVEC VKAESAQASAKMYEEMGKYQQMMEEKEKS YQEHVKQLTEKMERERAQLLEEQEKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKYMSHKLKI QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKYMSHKLKI QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKYMSHKLKI RFBRAPLVFSFEDVSEPKEKL RFBRAPLVFSFEDVSEPKEKL RFBRAPLVFSFEDVSEPKEKL RFBRAPLVFSFEDVSEPKEKL RFBRAPLVFSFEDVSEPKEKL RFBRAPLVFSFEDVSEPKEKL RFBRAPLVFSFEDVSEPKEKL RFBRAPLVFFVTMFRFTNTKNSDFFTBATEFT EGRFALALGGGYLHWGHFEMMRLTINRSM DPKNMFAINRVFAGRLVVFSFEDVSEPKEKL RFBRAPLVFMYTFERNATAMIGKVLSPYDL THKGKYWGKFYMFKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTITISDMKGSYFVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NITRGTREIGSALTRMCMRHISHEAKLRQFSS ALIDCLNFILQEQMEEWKNOQLGKDHAK EYKKARQEIKKRSSDTLKLQKKAKGRGDIQ PQLDSALQDVDNKYLLLEETKQAVRALIE ERGRFCTFISMLPFVUEELISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILLDLNFALSVSSH DSGFISQDAPGKSFSPMPFPAPNQRKKERE PDPNGGGFTTASGPPAAAEEAQRPSSM YQTPPSSFTTMSRKSSVCSSLNSVNSDSRS GSHSHISPSSITYRKSSNLAQQAFVRLSVSH DSGFISQDAPGKSSFSPMPFPAPNQRKEERE PDPNGGGFTTASGPPAAAEEAQRPSSM AGRGFGEKPSWSGFGSFGGTGRSWVKST RGHSSLLPPSQDFVAGLSVLRGTVDDRLNW AFNLTDLNKDGCITKEEMLDIMKSIDDMG KYYTYALAEAPREHVESFFQKMDRNKKDG VYYTYALAEAPREHVESFFQKMDRNKKDG KYYTYALAEAPREHVESFFQKMDRNKKDG WYTYTYALAEAPREHVESFFQKMDRNKKDG WYTYTYALAEAPREHVESFFQKMDRNKKDG WYTYTYALAEAPREHVESFFQKMDRNKKDG WYTYTYALAEAPREHVESFFQKMDRNKKDG WYTYTYALAEAPREHVESFFQKMDRNKKDG WYTYTYALAEAPREHVESFFQKMDRNKKDG WYTYTYALEEAPREHVESFFQKMDRNKKDG WYTYTYALAEAPREHVESFFQKMDRNKKDG WYTYTYALEEAPREHVESFFQKMDRNKKDG WYTYTYALAEAPREHVESFFQKMDRNKKDG WYTYTYTYDLNKNGGCITKEEMLDIMKSITDMG			1	Ì			
sequence nucleotide insertion	ì	l	l	l	t	sequence	
LMNKLAGKNKGFSLGSTVKSHTKGIWMWCV PHKKPEHTLVLLDTGGLGDVKKGDNQNS WIFTLAVLLSSTLVYNSMGTINQQAMDQLYY VTELTHRIRSKSSPDENEDSJADPVSFPPDFV WTLRDFSLDLEADGGPLTPDFYLEYSLKLTQ GTSQKDKNFNLPRLCIRKFPKKKCFVFDLPI HRRKLAQLEKLQDEELDPEFVQQVADFCSYI FSNSKTKTLSGGIKVNGPRLESLVLTYINAISR GDLPCMENAVLALAQIENSAAVQKAIAHYD QQMGQKQVQLPAETLGELLDLHRVSEREATEV YMKNSFKDVDHI-FQKKLAAQLDKKRDDFCK QNGRASDRCSALLQVFSLEEVKAGIVSK PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL QTYLKSKESVTDALQTDQLITEKEKEIEVEC VKAESAQSAKMVEEMQIKYQQMMEKEKS YQEHVKQLTEKMERERAQLLEEQEKTLTSKL QEQAVLKERCQGESTQHEIQKLQKTLKKK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TQLFAPLSGVSTRQLGSGAPLLTWVQETAG VAGGAPRRTTYTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVSFEDSUSPEKFKL RFIERAPLYPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAHWVPAPSFIVRLSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FILDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGIFOTIDSDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NIRGGTREIGSALTRMCMRHRISTAKLRGFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALGQVPNDKYLLDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSN VGTPPSSPSTTMSRKSSVCSSLNSVNSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSNS VGTPPSSPSTTMSRKSSVCSSLNSVNSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSNS VGTPPSSPSTTMSRKSSVCSSLNSVNSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSNS VGTPPSSPSTTMSRKSSVCSSLNSVNSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSNS VGTPPSSPSTTMSRKSSVCSSLNSVNSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSNS VGTPPSSPSTTMSRKSSVSSLNSVNSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSNS VGTPSSPSTTMSRKSSVSSLNSVNSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSNS VGTPSSPSTTMSRKSSVSSLNSVNSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSNS VGTPSSPSTTMSRKSSVSSSNSVNSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSN VGTPSSPSTTMSRKSSPSOGOTOGRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCTKEEMLDIMKSTYDMMG KYTTPALREEAPREHVEFSFQKMDRNKKDGV			ĺ				
PHPKKPEHTLVLLDTEGLGDVKKGDNONDS WIFTLAVILSSTLVYNSMGTINQQAMDQLYY VTELTHRIRSKSSPDENENEDSADFVSPFPDFV WTLADFSLDLEADGQPLTPEY-LEYSLKLTQ GTSQKDKNFNLPRLCIRKFPFKKCFYPDLPI HRRKLAQLEKLQDEGLDPEFVQQVADFCSYI FSNSKTKILSGGIKVNGPRLESLVLTYINAISR GDLFCMENAVLALAQLEKKAQLEKKADPCK QQMGQKVQLPAETLQELLDLHRVSEREATEV YMKNSFKDVDHLFGKAAQLDKKRDDFCK QNQEASSDRCSALLQVIFSPLEEEVKAGIYSK PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL QTYLKSKESVTDAILQTDQILTEKEKEIEVEC VKAESAQASAKMVEEMQIKYQQMMEEKEKS YQEHVKQLTEKMERERAQLLEEQEKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TQLPAPLSGGSTPLQNEIQKLQKTLKK KTKRYMSHKLKI DFLANGARPRITYTMWRLLARASAPLLRVPLS SDWALLFBASAGVKTLDYSFEDVSIPEKFKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWCHFERMRTINTASM DPKNMFAINEVPAPFFTNRSVGHRMGGGK GADHYVTFVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLFFAAKAVGTLEKMRKDOEE RERNIQNPWTTERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLPQTITISDMKGSYFVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTEGGSALTEMGRRSQLKALRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKRSDTLKLQKKAKGRGDIO QPQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGFCTFISMLRPVILEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSENLAQQAPVLLSVSSSH SGHSHSPSHYNRSSNLAQQAPVLLSVSSSH SGHSHSPSSHYNRSSNLAQQAPVLLSVSSSH SGHSHSPSSHYNRSSNLAQQAPVLLSVSSSH DSGFISQDAFQSKSPSRMPPEAPNQRKEKRE PDPNGGGTTTASQPPAAAEEAQRRSM AFNLYDLNKDGCTIKEEMLDIMKSTYDMMG KYTYPALREEAFREHVEFFQKMDRNKDGV				ļ	sequence		
WIFTLAYLLSSTLVYNSMGTINQQAMDQLYY VTELTHRIRSKSSPDENENEDSADFVSFFPDFV WTIRADFSLDLEADGQPLTPDEYLEYSLKLTQ GTSQKDKNFNI.PRLCIRKFFFKKK.CFVFDLPI HRRILAQLEKLQDEELDPEVQQVADFCSYI FSNSKTKTLSGGIKVNGPRLESLVLTYINAISR GDLPCMENAVLALAQIENSAAVQKAIAHYD QQMQQKVQLPAETLQELDLHRYSEREATEV YMKNSFKDVDHLFQKKLAQLDKKRDDFCK QNQEASDRCSALLQVDLEKKYYEEPRKGIQAEEL QTYLKSKESVTDALI.ODLTEKEKEEIEVEC VKAESAQASAKMVEEMQIKYQQMMEEKEKS YQEHVKQLTEKMERERAQLLEGGKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TQLPAPLSGVISHCKEYC VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVSFEDVSIPEKFKL RFIERAPLYPKVRREPKNLSDIRGFSTEATEFT EGNFALLALGGGYLHWGHFEMMRLTINRSM DPKNMFAINWPAPFRNLSDIRGFSTEATEFT EGNFALLALGGGYLHWGHFEMMRLTINRSM DPKNMFAINWPAPFRNLSDIRGFSTEATEFT EGNFALLALGGGYLHWGHFEMMRLTINRSM GADHYVTPVKAGRLVVEMGGGCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RENNIQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFOTISDMKGSYFVWEDFINK AGKLQSQLRTTVVAAAAFLDAFGKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKVANQLDKDCHAK EYKKARQEIKKKSSDTLKLQKKAKGRGDIQ PQLDSALGQVDNDKYLLLEETERQAVKKALE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSENDLAQGRAVELSEVSSM YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRCSSVCSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRCSSVCSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRCSSVCSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRSSVCSSLNSVNSSDSRSS GSHSHSPSSHYKYRSSNLAQQAPVLSSVSSM YQTPPSSPSTTMSRSSPQTTQRSWVKST RGHS				ì			
VTELTHRIRSKSSPDENENEDSADFVSFFPDFV WTI.RDFSLDLEADGQPLTPDEYLEYSLKLTQ GTSQKDKNFN.PRLCIRKFFPKKK.CFVFDLPI HRRKLAQLEKLQDEELDPEFYQQVADFCSYI FSNSKTKTLSGGIKVNOFRLESLV.TYINAISR GDLPCMENAVLALAQIENSAAVQKAIAHYD QQMGQKVQLPAETLQELDLHRVSEREATEV YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK QNQEASSDRCSALLQVIFSLEEVVAGIYSK PGGYCLFJOKLQDLEKYYPEPRKGIQAEEL QTYLKSKESVTDALQTIDQILTEKEKEIEVEC VKAESAQASAKMYEEMQIKYQQMMEEKEKS YQEHVKQLTEKMERERAQLLEEQEKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTIPVTMWRLLAARSAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREFKNLSDIRGPSTEATEFT EGNFALLALGGGYLHWGHFEMMRLTINGSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVEEMGGCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRDQEE RERNINQNPWTFERIATANNLGIRKVLSPYDL THKGKYWGKPYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLGFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHSIESEAKLRQFSS ALIDCLNPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVKKALE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSEQVILDLKGSDYSW YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRKEKRE PDPNGGGFTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGGFGFVSWSGFSPGGTGGTGRSWWKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCTTKEEMLIDMKSTYDDMMG KYTYPALREEAPREHVESFFQKMDRNKDOV	1	ł		1			1
WTLRDFSLDLEADGOPLTPDEYLEYSLKLTQ GTSQKDKNFNLPRLCIRKFPFKKCFVFDLPI HRRKLAQLEKLQDEELDPEFVQQVADFCSYI FSNSKTKTLSGGIKVNOPRLESLVLTYINAISR GDLPCMENAVLALQENSAAVQKAIAHYD QQMGQKVQLPAETLQELLDLHRVSEREATEV YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK QNQEASSDRCSALLQVIFSPLEEVKAGIYSK PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL QTYLKSKESVTDAILQVIFSPLEEVKAGIYSK PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL QTYLKSKESVTDAILQVIFSPLEEVKAGIYSK PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL QTYLKSKESVTDAILQVIFSPLEEVKAGIYSK KAESAQAASAKNYEMQIKYQQMMEEKEKS YQEHYKQLTEKMERERAQLLEEQEKTLTSKL QEQARVLKERCQGETQLONEIQKLQKTLKK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TQLPAPLISGYLSRLQLGSGAPLLTWVQETAG VAGGAPRRTIPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFALLALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITKSVGHRMGGGK GADHYVTPVXAGRUVEMGGRCEFEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNONPWTFERIATAMIGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEERINGEITHLQTITSE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSINSVNSSDSRSS GSHSHSPSSHYRYSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGFTTASGPPAAAEEAQRRSSM VGTPPSSPSTTMSRKSSVCSSINSVNSSDSRSS GSHSHSPSSHYRYSSNLAQQAPVALSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGFTTASGPPAAAEEAQRRSSW YGTPPSSPSTTMSRKSSVCSSINSVNSSDSRSS GSHSHSPSSHYRYSSNLAQQAPVALSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGFTTASGPPAAAEEAQRRSSW KST RGHSSLLPPSQDPVAGLSVLRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMORNKDOV							
GTSQKDKNFNLPRLCIRKFFPKKKCFVPDLP HRRKLAQLEKLQDEELDPEFVQQVADFCSYI FSNSKTKTLSGGIKVNOFRLESLVLTYINAISR GDLPCMENAVLALLQEINSAAVQKAIAHYD QOMGOKVQLPAFLQELDJHRYSEREATEV YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK QNQEASSDRCSALLQVIFSIEEVKAGIYSK PGGYCLFQKLQDLEKKYYEEPRKGIQAEEL QTYLKSKESVTDAILQTIQLITEKEKEIEVEC VKAESAQASAKMYEEMQIK YQQMMEEKEKS YQEHVKQLTEKMERERAQLLEEQEKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TOLPAPLSGYLSKLQLGSGAPLLTWVQETAG VAGGAPRRTPVTWWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPYPSSEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNPAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDARQKVADMAT NTRGGTREIGSALTROKRHRSIESAKLRQFSS ALIDCLINPLQEMEWKKVANQLDKDHAK EYKKARQEIKKKSSEQVILDLKGKAKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTITSE DLKSLTMDPHKLIPSESQVILDLKGSDYSW YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQKKSPSPMPPEAPNQRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGGFGKPVSWSSGPSGGTGTGRSWVKST RGHSSLLPPSQDPVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLIDMKSIYDMMG	1	1	ļ			,	
HRRKLAQLEKLQDEELDPEFVQQVADFCSYI FSNSKTKIT.SGIKVNGPRLESLVLTYINAISR GDLPCMENAVLALAQIENSAAVQKAIAHYD QQMGQKVQLPAETI.QELLDLHRVSEREATEV YMKNSFKDVDHLFOKKLAAQLDKKRDDPCK QNQEASSDRCSALLQVIFSPLEEEVKAGIYSK PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL QTYLKSKESVTDAILQTDQLITEKKEEIEVEC VKAESAQASAKMVEEMQIKYQQMMEEKEKS YQEHVKQLTEKMERERAQLLEEQEKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHLKI 1348 2698 A 10404 5 892 TQLPAPLSGVLSERQLEQEKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHLKI TQLPAPLSGVLSRCQESTQLQNEIQKLQKTLKK KTKRYMSHLKI RFIERAPLYPKVREPFKNLSDIRGPSTEATEFT EGNFALLALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAFFKPITRKSVGHRMGGK RFIERAPLYPKVREPFKNLSDIRGPSTEATEFT EGNFALLALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAFFKPITRKSVGHRMGGK GAIDHYVTIPVKAGRLVVEMGGRCEFFEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNONPWTEATAANMLGIRKVLSPYDL THKGKYWGKFYMPKRV TKKGKYGKFYMPKRV TKKGKYWGKFYMPKRV TKKGKYWGKFYMPKRV TKKGKYWGKFYMPKRV TKKGKYWGKFYMPKRV TKKGKYMGKYTHKKFK TKKGKYMGKYTHKKFK TKKGKYMGKYTHKKFK TKKGKYMGKYTHKKFK TKKGKYMGKYTHKKFK TKKGKYMGKYTHKKFK TKKGKYMGKYTHKKFK TKKGKYMGKYTHKKFK TKKGKYMGKYTHKKKGCTTKEMLT TKKGKYMGKYTHKKFK TKKGKYMGKYTHKKFK TKKGKYMGKYTHKKGCTTKEMLT TKKGKYMGKTHK TKKGKYMGKTHK TKKGKTHK TKKGKTHK TKKGKTHK TKKGKTHK TKKGKTHK TKKGKTHK TKKGKTHK TKKGKTHKT	l						
FSNSKTKTLSGGIKVNGPRLESLVLTYINAISR GDLPCMENAVLALAQIENSAAVQKAIAHYD QQMGQKVQLPAFTLQELLDLIRVSEREATEV YMKNSFKDVDHLFQKKLAQLDKKRDDFCK QNQEASSDRCSALLQVIFSPLEEVKAGIYSK PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL QTYLKSKESVTDAILQTDQILTEKEKEIEVEC VKAESAQASAKVEEMQIKYQOMMEEKEKS YQEHVKQLTEKMERERAQLLEEQEKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI. 1348 2698 A 10404 5 892 TQLPAPLSGVLSRQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVSFEDVSIPEKPKL RFIERAPLYFVRWREPKNLSDIRGFSTEATEFT EGNFALLALGGGYLHWGHFEMMRLTINRSM DPKNMFALWRVPAPFKPITRKSVGHRMGGGK GADHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMGIRKVLSPYDL THKGKYWGKPYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTIMCMRHISIEAKLRQFSS ALIDCLINPLQEGUMEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGETHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDRSSS GSHSSHSPSSHYNYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPHQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGRFVSWSSGPGSPGOTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV]	i			UDDKI ACI EKI ODDEI DDETVOOMADDOOM
GDLPCMENAVLALAQIENSAAVQKAIAHYD QQMGQKVQLPAETLQELLDLHRVSEREATEV YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK QNQEASSDRCSALLQVIFSPLEEWKAGIYSK PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL QTYLKSKESVTDAILQTDQILTEKEKEIEVEC VKAESAQASAKMVEEMQIKYQQMMEEKEKS YQEHVKQLTEKMERERAQLLEEQEKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPYTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GADHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNONPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKKV 1349 2699 A 10409 59 1184 LRRNCSALGGLTTUSDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRRTSIEAKLRQFSS ALIDCLINPLQEQMEWKKVANQLDKDHAK EYKKARQEIKKRSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDRSS GSHSHSPSSHYXYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRREKRE PDPNGGGPTTASQPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGFGKPVSWSSGFGSPGGTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFGKMDRNKDGV	1	•					FORESTET SOCIEDATION FOR THE TYPE A TOP
QQMGQKVQLPAETLÖELDLHRVSEREATEV YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK QNQEASSDRCSALLQVIFSPLEEDVKAGIYSK PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL QTYLKSKESVTDAILQTDQILTEKEKEEIEVEC VKAESAQASAKMYEEMQIKYQQMMEEKEKS YQEHVKQLTEKMERERAQLLEEQEKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TQLPAPLSGVISRLQLGSGAPLLTWVQETAG VAGGAPRRTPYTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFALLALGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKNMKLOGE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMFKV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALDDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSISPSSTHYSYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRREKERE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGFGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVLRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	}			}			
YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK QNQEASSDRCSALLQVIFSPLEEEVKAGIYSK PGGYCLFIQKLQDLEKKYYEBPRKGIQAEBIL QTYLKSKESVTDAILQTDQILTEKEKEIEVEC VKAESAQASAKMYEMQIKYQQMMEEKEKS YOEHYKQLTEKMERRAQLLEEQEKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI QEQARVLSELQLGGGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVFSFEDVSIPEKPKL RFIERAPLVPKVREFEKNLSDINGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHEMGGGK GADHYVTPVAGRCTVEVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV THKGKYWGKFYMPKRV AGKLQSQLRTTVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTIKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRKEKRE PDPNGGGPTTASGDPAAAEEAQRPSM TGHSSLLPPSQDPAGAGEAQRPSM AGRGGGGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDPAGSLSVLRGTVODRLNW AFNLYDLNKGGCTKEEMLDIMKSIYDDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	1	1	Ì				
QNQEASSDRCSALLQVIFSPLEEEVKAGIYSK PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEII QTYLKSKESYTDAILQTDQILTEKEKEIEVEC VKAESAQASAKMVEEMQIKYQQMMEEKEKS YQEHVKQLTEKMERERAQLLEEQEKTLTSKL QEQARVLKERCQESTQLQNEIQKLQKTLKK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPYTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPFKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAILWVPAPFKPITKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDARQSKSPSMPPEAPNQRKKEKRE PPDPNGGGPTTASGPPAAAEEAQPRSM 1350 2700 A 10410 511 958 AGRGGFGKPVSWSSGFGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV		ŀ					
PGGVCLFIQKLQDLEKKYYEEPRKGIQAEEIL QTYLKSKESVTDAILQTDQILTEKEKEIEVEC VKAESAQASAKMVEEMQIKYQQMMEEKEKS YQEHVKQLTEKMERERAQLLEEQEKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSSEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILAGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITTRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANNLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLNPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGFCTFISMLRPVIEEFISMLGEITHLQTISE DLKSLTMDPHKLPSSEQVILDLKGSDVSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSH DSGFISQDAFQSKSSPSMPPEAPNQRRKEKRE PPDPNGGGPTTASGPPAAAEEAQPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPSGGTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV			İ				UNUEVSCUDGE VI I UNIESDI EEEMA VOINSIA
QTYLKSKESVTDAILQTDQILTEKEKEEVEC VKAESAQASAKMYEEMQIKYQQMMEEKEKS YQEHVKQLTEKMERERAQLLEEQEKTLTSKL QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARGUKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRVRSSNLAQQAPVRLSSVSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGTTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESSFQKMDRNKDGV	1	}					
VKAESAQASAKMVEĖMQIKYQQMMEEKEKS YQEHYKQLITEKMERERAQLIEEQEKTI.TSKL QEQARVI.KERCOGESTQLQNEIQKIQKTI.KK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TQLPAPLSGVLSRLQLGSGAPILITWVQETAG VAGGAPRRITYTYMWRLLARASAPILARVILS DSWALLIPASAGVKTILIPYPSFEDVSIPEKPKI. RFIERAPLVPKVRREPKNI.SDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRI.TINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVIPVKAGRI.VVEMGGRCEFFEVQG FILDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTTERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NITRGGTREIGSALTRMCMRIRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGTTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGFGKPVSWSSGPSGTQTQRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	1						
1348 2698 A 10404 5 892 TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RIERRAPLVPKVREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAILWRVPAFFKPITRKSVGHRMGGGK GADHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRINCSALGGIFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVYKALIE ERGRFCTFISMLRPVIEELISMLGEITHLQTISE DLKSSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGGGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVLRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	1						
QEQARVLKERCQGESTQLQNEIQKLQKTLKK KTKRYMSHKLKI 1348 2698 A 10404 5 892 TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPYTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFALALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GADHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGEFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKRSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYPRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVODRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV							
STERRYMSHKLKI 1348 2698 A 10404 5 892 TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLLYDLINKDGCTIKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV		i		ł			
1348 2698 A 10404 5 892 TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTILKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILDHKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV							
VAGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNINQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKPYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGGGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCTTKEEMLDIMKSTYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	1348	2698	A	10404	5	892	
DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNI.SDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRSSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FILDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSISHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGEKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV							
RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FILDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPPNGGGFTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGGGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCTTKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	İ		i				
DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPSGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMI.DIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	1						
GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	1			}			EGNFAILALGGGYLHWGHFEMMRLTINRSM
FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	i						DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK
RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	1						GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG
THKGKYWGKFYMPKRV 1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV							
1349 2699 A 10409 59 1184 LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRKEKRE PDPPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	l	}					
AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRKEKRE PDPPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV		2.455					
NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRKEKRE PDPPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	1349	2699	Α	10409	59	1184	LRRNCSALGGLFQTIISDMKGSYPVWEDFINK
ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	1	İ					AGKLQSQLRTTVVAAAAFLDAFQKVADMAT
EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	l						
PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	!						
ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	j	1					
DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV							
YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV							
GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV							AULDESCELAM LIVER 399EA AITHER AS 12 A L
DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	[[
PDPNGGGPTTASGPPAAAEEAQRPRSM 1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV		[
1350 2700 A 10410 511 958 AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV							
RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV	1350	2700	A	10410	511	958	
AFNLYDLNKDGCITKEEMLDIMKSIYDMMG KYTYPALREEAPREHVESFFQKMDRNKDGV							
KYTYPALREEAPREHVESFFQKMDRNKDGV	1						
				1			I ACINE I DENNEDUCI I KERIVII I JIMIK ALT I JIMIVII I I I
	}						

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, and complementary sequences thereof.

- 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
- 4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- 5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
- 6. A vector comprising the polynucleotide of claim 1.
- 7. An expression vector comprising the polynucleotide of claim 1.
- 8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
- 9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
- 10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
 - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
 - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-1350.
- 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and

- b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
- 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that annual to the polynucleotide of claim 1 under such conditions;
- b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
- 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
- b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
- 17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 19. A method of producing the polypeptide of claim 10, comprising,
- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-1350, under conditions sufficient to express the polypeptide in said cell; and
 - b) isolating the polypeptide from the cell culture or cells of step (a).
- 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1351-2700, the mature protein portion thereof, or the active domain thereof.
- 21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
- 22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-1350.
- 23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
- 24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
- 25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
- 26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

Pages 340 to 1963 of this application contain amino acid sequence listings. They can be obtained at the address given below.

Les pages 340 to 1963 de cette demande contiennent des listages des séquences d'acides aminés. Elles peuvent être obtenues à l'adresse indiquée ci-dessous.

World Intellectual Property Organization 34, chemin des Colombettes CH-1211 Genève 20

(19) World Intellectual Property Organization International Bureau





(43) International Publication Date 9 August 2001 (09.08.2001)

PCT

(10) International Publication Number WO 01/57188 A3

(51) International Patent Classification⁷: C07K 5/00, A61K 39/395, C12Q 1/68

C07H 21/04.

(21) International Application Number: PCT/US01/03800

(22) International Filing Date: 5 February 2001 (05.02.2001)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

09/496,914 09/560,875 3 February 2000 (03.02.2000) US 27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 09/496,914 (CIP)
Filed on 3 February 2000 (03.02.2000)
US 09/560.875 (CIP)

Filed on

27 April 2000 (27.04.2000)

(71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court. San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14. San Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

(74) Agent: ELRIFI, Ivor, R.: Mintz, Levin, Coh, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

- (81) Designated States (national): AE. AG. AL. AM. AT. AU. AZ. BA. BB. BG, BR. BY. BZ. CA. CH. CN. CR. CU. CZ. DE, DK. DM. DZ. EE. ES. FI. GB. GD. GE. GH. GM. HR. HU, ID, IL, IN, IS, JP. KE. KG. KP. KR. KZ. LC. LK. LR. LS. LT. LU. LV. MA. MD. MG. MK. MN. MW, MX. MZ. NO, NZ. PL. PT. RO. RU, SD. SE. SG, SI, SK. SL. TJ. TM. TR. TT. TZ. UA. UG. US. UZ, VN, YU. ZA. ZW.
- (84) Designated States (regional): ARIPO patent (GH. GM. KE, LS, MW. MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau
- (88) Date of publication of the international search report: 28 February 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

/57188 A

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

INTERNATIONAL SEARCH REPORT

International application No. PCT/US01/03800

A. CLASSIFICATION OF SUBJECT MATTER									
IPC(7) :C07H 21/0+; C07K 5/00; A61K 39/395; C12Q 1/68									
US CL :586/23.1; 530/300; +2+/130.1; +35/6									
According to International Patent Classification (IPC) or to both national classification and IPC									
B. FIELDS SEARCHED									
Minimum documentation searched (classification system followed	by classification symbols)								
U.S. : 536/23.1; 530/300; +2+/130.1; +35/6									
Documentation searched other than minimum documentation to searched	Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched								
Electronic data base consulted during the international search (n STN, medline search terms: EST, expressed sequence tags									
C. DOCUMENTS CONSIDERED TO BE RELEVANT									
Category* Citation of document, with indication, where ap	propriate, of the relevant passages Relevant to claim No.								
Map Based on Naturally Occurring	action of document, with indication, where appropriate, of the relevant publication								
Further documents are listed in the continuation of Box (C. See patent family annex.								
* Special categories of cited documents:	"T" later document published after the international filing date or priority								
"A" document defining the general state of the art which is not considered	date and not in conflict with the application but cited to understand the principle or theory underlying the invention								
to be of particular relevance	"X" document of particular relevance, the claimed invention cannot be								
"L" earlies document published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is	considered novel or cannot be considered to involve an inventive step when the document is taken alone								
cited to establish the publication date of another citation or other	"Y" document of particular relevance: the claimed invention cannot be								
special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other	considered to involve an inventive step when the document is combined with one or more other such documents, such combination being								
O" document referring to an oral disclosure, use, exhibition or other with one or more other such documents, such combination being means obvious to a person skilled in the art									
"P" document published prior to the international filing date but later than the priority date claimed	"A" document member of the same patent family								
Date of the actual completion of the international search	Date of mailing of the international search report								
so SEPTEMBER 2001	16 NOV 2001								
Name and mailing address of the ISA/US	Authorized officer JEHANNE SOUAYA Telephone No. (703) 308-1235								
Commissioner of Patents and Trademarks Box PCT	The 1 11/16/1								
Washington, D.C. 20231	JEHANNE SOUAYA								
Facsimile No. (703) 305-3230	Telephone No. (703) 308-1235								

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:
□ BLACK BORDERS
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
☐ FADED TEXT OR DRAWING
☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
☐ SKEWED/SLANTED IMAGES
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
☐ GRAY SCALE DOCUMENTS
☐ LINES OR MARKS ON ORIGINAL DOCUMENT
\square REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

IMAGES ARE BEST AVAILABLE COPY.

OTHER:

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.